STATISTICAL METHODS FOR SIGNAL PROCESSING

Alfred O. Hero

August 25, 2008

This set of notes is the primary source material for the course EECS564 "Estimation, filtering and detection" used over the period 1999-2007 at the University of Michigan Ann Arbor. The author can be reached at
Dept. EECS, University of Michigan, Ann Arbor, MI 48109-2122
Tel: 734-763-0564.
email hero@eecs.umich.edu;
`http://www.eecs.umich.edu/~hero/`.

Contents

1 INTRODUCTION

1.1 STATISTICAL SIGNAL PROCESSING

Many engineering applications require extraction of a signal or parameter of interest from degraded measurements. To accomplish this it is often useful to deploy fine-grained statistical models; diverse sensors which acquire extra spatial, temporal, or polarization information; or multi-dimensional signal representations, e.g. time-frequency or time-scale. When applied in combination these approaches can be used to develop highly sensitive signal estimation, detection, or tracking algorithms which can exploit small but persistent differences between signals, interferences, and noise. Conversely, these approaches can be used to develop algorithms to identify a channel or system producing a signal in additive noise and interference, even when the channel input is unknown but has known statistical properties.

Broadly stated, statistical signal processing is concerned with the reliable estimation, detection and classification of signals which are subject to random fluctuations. Statistical signal processing has its roots in probability theory, mathematical statistics and, more recently, systems theory and statistical communications theory. The practice of statistical signal processing involves: (1) description of a mathematical and statistical model for measured data, including models for sensor, signal, and noise; (2) careful statistical analysis of the fundamental limitations of the data including deriving benchmarks on performance, e.g. the Cramèr-Rao, Ziv-Zakai, Barankin, Rate Distortion, Chernov, or other lower bounds on average estimator/detector error; (3) development of mathematically optimal or suboptimal estimation/detection algorithms; (4) asymptotic analysis of error performance establishing that the proposed algorithm comes close to reaching a benchmark derived in (2); (5) simulations or experiments which compare algorithm performance to the lower bound and to other competing algorithms. Depending on the specific application, the algorithm may also have to be adaptive to changing signal and noise environments. This requires incorporating flexible statistical models, implementing low-complexity real-time estimation and filtering algorithms, and on-line performance monitoring.

1.2 PERSPECTIVE ADOPTED IN THIS BOOK

This book is at the interface between mathematical statistics and signal processing. The idea for the book arose in 1986 when I was preparing notes for the engineering course on detection, estimation and filtering at the University of Michigan. There were then no textbooks available which provided a firm background on relevant aspects of mathematical statistics and multivariate analysis. These fields of statistics formed the backbone of this engineering field in the 1940's 50's and 60's when statistical communication theory was first being developed. However, more recent textbooks have downplayed the important role of statistics in signal processing in order to accommodate coverage of technological issues of implementation and data acquisition for specific engineering applications such as radar, sonar, and communications. The result is that students finishing the course would have a good notion of how to solve focussed problems in these applications but would find it difficult either to extend the theory to a moderately different problem or to apply the considerable power and generality of mathematical statistics to other applications areas.

The technological viewpoint currently in vogue is certainly a useful one; it provides an essential engineering backdrop to the subject which helps motivate the engineering students. However, the disadvantage is that such a viewpoint can produce a disjointed presentation of the component parts of statistical signal processing making it difficult to appreciate the commonalities between detection, classification, estimation, filtering, pattern recognition, confidence intervals and other useful tools. These commonalities are difficult to appreciate without adopting a proper statistical perspective. This book strives to provide this perspective by more thoroughly covering elements of mathematical statistics than other statistical signal processing textbooks. In particular we cover point estimation, interval estimation, hypothesis testing, time series, and multivariate analysis. In adopting a strong statistical perspective the book provides a unique viewpoint on the subject which permits unification of many areas of statistical signal processing which are otherwise difficult to treat in a single textbook.

The book is organized into chapters listed in the attached table of contents. After a quick review of matrix algebra, systems theory, and probability, the book opens with chapters on fundamentals of mathematical statistics, point estimation, hypothesis testing, and interval estimation in the standard context of independent identically distributed observations. Specific topics in these chapters include: least squares techniques; likelihood ratio tests of hypotheses; e.g. testing for whiteness, independence, in single and multi-channel populations of measurements. These chapters provide the conceptual backbone for the rest of the book. Each subtopic is introduced with a set of one or two examples for illustration. Many of the topics here can be found in other graduate textbooks on the subject, e.g. those by Van Trees, Kay, and Srinath *etal.* However, the coverage here is broader with more depth and mathematical detail which is necessary for the sequel of the textbook. For example in the section on hypothesis testing and interval estimation the full theory of sampling distributions is used to derive the form and null distribution of the standard statistical tests of shift in mean, variance and correlation in a Normal sample.

The second part of the text extends the theory in the previous chapters to non i.i.d. sampled Gaussian waveforms. This group contains applications of detection and estimation theory to single and multiple channels. As before, special emphasis is placed on the sampling distributions of the decision statistics. This group starts with offline methods; least squares and Wiener filtering; and culminates in a compact introduction of on-line Kalman filtering methods. A feature not found in other treatments is the separation principle of detection and estimation which is made explicit via Kalman and Wiener filter implementations of the generalized likelihood ratio test for model selection, reducing to a whiteness test of each the innovations produced by a bank of Kalman filters. The book then turns to a set of concrete applications areas arising in radar, communications, acoustic and radar signal processing, imaging, and other areas of signal processing. Topics include: testing for independence; parametric and non-parametric testing of a sample distribution; extensions to complex valued and continuous time observations; optimal coherent and incoherent receivers for digital and analog communications;

A future revision will contain chapters on performance analysis, including asymptotic analysis and upper/lower bounds on estimators and detector performance; non-parametric and semiparametric methods of estimation; iterative implementation of estimators and detectors (Monte Carlo Markov Chain simulation and the EM algorithm); classification, clustering, and sequential design of experiments. It may also have chapters on applications areas including: testing of binary Markov sequences and applications to internet traffic monitoring; spatio-temporal signal processing with multi-sensor sensor arrays; CFAR (constant false alarm rate) detection strategies for Electro-optical (EO) and Synthetic Aperture Radar (SAR) imaging; and channel equalization.

1.2.1 PREREQUISITES

Readers are expected to possess a background in basic probability and random processes at the level of Stark&Woods [68], Ross [59] or Papoulis [54], exposure to undergraduate vector and matrix algebra at the level of Noble and Daniel [52] or Shilov [64] , and basic undergraduate course on signals and systems at the level of Oppenheim and Willsky [53]. These notes have evolved as they have been used to teach a first year graduate level course (42 hours) in the Department of Electrical Engineering and Computer Science at the University of Michigan from 1997 to 2008 and a one week short course (40 hours) given at EG&G in Las Vegas in 1998.

The author would like to thank Hyung Soo Kim, Robby Gupta, and Mustafa Demirci for their help with drafting the figures for these notes. He would also like to thank the numerous students at UM whose comments led to an improvement of the presentation. Special thanks goes to Clayton Scott of the University of Michigan, Raviv Raich of Oregon State University and Aaron Lanterman of Georgia Tech who provided detailed comments and suggestions for improvement of earlier versions of these notes. **End of chapter**

2 NOTATION, MATRIX ALGEBRA, SIGNALS AND SYSTEMS

Keywords: vector and matrix operations, matrix inverse identities, linear systems, transforms, convolution, correlation.

Before launching into statistical signal processing we need to set the stage by defining our notation. We then briefly review some elementary concepts in linear algebra and signals and systems. At the end of the chapter you will find some useful references for this review material.

2.1 NOTATION

We attempt to stick with widespread notational conventions in this text. However inevitably exceptions must sometimes be made for clarity.

In general upper case letters, e.g. X, Y, Z, from the end of the alphabet denote random variables, i.e. functions on a sample space, and their lower case versions, e.g. x, denote realizations, i.e. evaluations of these functions at a sample point, of these random variables. We reserve lower case letters from the beginning of the alphabet, e.g. a, b, c, for constants and lower case letters in the middle of the alphabet, e.g. i, j, k, l, m, n, for integer variables. Script and caligraphic characters, e.g. $\mathcal{S}$, $\mathcal{I}$, Θ, and $\mathcal{X}$, are used to denote sets of values. Exceptions are caligraphic upper case letters which denote standard probability distributions, e.g. Gaussian, Cauchy, and Student-t distributions $\mathcal{N}(x), \mathcal{C}(v), \mathcal{T}(t)$, respectively, and script notation for power spectral density $\mathcal{P}_x$. Vector valued quantities, e.g. $\underline{x}, \underline{X}$, are denoted with an underscore and matrices, e.g. $\mathbf{A}$, are bold upper case letters from the beginning of the alphabet. An exception is the matrix $\mathbf{R}$ which we use for the covariance matrix of a random vector. The elements of an $m \times n$ matrix $\mathbf{A}$ are denoted generically $\{a_{ij}\}_{i,j=1}^{m,n}$ and we also write $\mathbf{A} = (a_{ij})_{i,j=1}^{m,n}$ when we need to spell out the entries explicitly.

The letter f is reserved for a probability density function and p is reserved for a probability mass function. Finally in many cases we deal with functions of two or more variables, e.g. the density function $f(x;\theta)$ of a random variable X parameterized by a parameter θ. We use subscripts to emphasize that we are fixing one of the variables, e.g. $f_\theta(x)$ denotes the density function over x in a sample space $\mathcal{X} \subset \mathbb{R}$ for a fixed θ in a parameter space Θ. However, when dealing with multivariate densities for clarity we will prefer to explicitly subscript with the appropriate ordering of the random variables, e.g. $f_{X,Y}(x, y; \theta)$ or $f_{X|Y}(x|y; \theta)$.

2.2 VECTOR AND MATRIX BACKGROUND

2.2.1 ROW AND COLUMN VECTORS

A vector is an ordered list of n values:

$$\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

which resides in $\mathbb{R}^n$.

Convention: in this course $\underline{x}$ is (almost) always a column vector. Its transpose is the row vector

$$\underline{x}^T = [\ x_1 \ \cdots \ x_n\ ]$$

When the elements $x_i = u + jv$ are complex (u, v real valued, $j = \sqrt{-1}$) the Hermitian transpose is defined as

$$\underline{x}^H = [\ x_1^* \ \cdots \ x_n^*\ ]$$

where $x_i^* = u - jv$ is the complex conjugate of x_i.

Some common vectors we will see are the vector of all ones and the j-th elementary vector, which is the j-th column of the identity matrix:

$$\underline{1} = [1, \ldots, 1]^T, \qquad \underline{e}_j = [0, \ldots, 0, \underbrace{1}_{j-th}, 0, \ldots 0]^T$$

2.2.2 VECTOR/VECTOR MULTIPLICATION

For 2 vectors $\underline{x}$ and $\underline{y}$ with the same number n of entries, their "inner product" is the scalar

$$\underline{x}^T \underline{y} = \sum_{i=1}^n x_i y_i$$

The 2-norm $\|\underline{x}\|_2$ of a vector $\underline{x}$ is its length and it is defined as (we drop the norm subscript when there is no risk of confusion)

$$\|\underline{x}\| = \sqrt{\underline{x}^T \underline{x}} = \sqrt{\sum_{i=1}^n x_i^2}.$$

For 2 vectors $\underline{x}$ and $\underline{y}$ of possibly different lengths n, m their "outer product" is the $n \times m$ matrix

$$\begin{aligned}
\underline{x}\underline{y}^T &= (x_i y_j)_{i,j=1}^{n,m} \\
&= [\underline{x} y_1, \ldots, \underline{x} y_m] \\
&= \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_m \\ \vdots & \ddots & \vdots \\ x_n y_1 & \cdots & x_n y_m \end{bmatrix}
\end{aligned}$$

2.3 ORTHOGONAL VECTORS

If $\underline{x}^T \underline{y} = 0$ then $\underline{x}$ and $\underline{y}$ are said to be orthogonal. If in addition the lengths of $\underline{x}$ and $\underline{y}$ are equal to one, $\|\underline{x}\| = 1$ and $\|\underline{y}\| = 1$, then $\underline{x}$ and $\underline{y}$ are said to be orthonormal vectors.

2.3.1 VECTOR/MATRIX MULTIPLICATION

Let $\mathbf{A}$ be an $m \times n$ matrix with columns $\underline{a}_{*1}, \ldots, \underline{a}_{*n}$ and $\underline{x}$ be any n-element vector.

The (compatible) product $\mathbf{A}\underline{x}$ is a (column) vector composed of linear combinations of the columns of $\mathbf{A}$

$$\mathbf{A}\underline{x} = \sum_{j=1}^{n} x_j \, \underline{a}_{*j}$$

For $\underline{y}$ an m-element vector the product $\underline{y}^T\mathbf{A}$ is a (row) vector composed of linear combinations of the rows of $\mathbf{A}$

$$\underline{y}^T\mathbf{A} = \sum_{i=1}^{m} y_i \, \underline{a}_{i*}.$$

2.3.2 THE LINEAR SPAN OF A SET OF VECTORS

Let $\underline{x}_1, \ldots, \underline{x}_n$ be a set of p dimensional (column) vectors and construct the $p \times n$ matrix

$$\mathbf{X} = [\underline{x}_1, \ldots, \underline{x}_n].$$

Let $\underline{a} = [a_1, \ldots, a_n]^T$ be a vector of coefficients. Then $\underline{y} = \sum_{i=1}^{n} a_i \underline{x}_i = \mathbf{X}\underline{a}$ is another p dimensional vector that is a linear combination of the columns of $\mathbf{X}$. The linear span of the vectors $\underline{x}_1, \ldots, \underline{x}_n$, equivalently, the column space or range of $\mathbf{X}$, is defined as the subspace of $\mathbb{R}^p$ that contains all such linear combinations:

$$\text{span}\{\underline{x}_1, \ldots, \underline{x}_n\} = \{\underline{y} : \underline{y} = \mathbf{X}\underline{a}, \ \underline{a} \in \mathbb{R}^n\}.$$

In other words, when we allow $\underline{a}$ to sweep over its entire domain $\mathbb{R}^n$, $\underline{y}$ sweeps over the linear span of $\underline{x}_1, \ldots, \underline{x}_n$.

2.3.3 RANK OF A MATRIX

The (column) rank of a matrix $\mathbf{A}$ is equal to the number its columns which are linearly independent. The dimension of the column space of a rank p matrix $\mathbf{A}$ is equal to p.

If $\mathbf{A}$ has full rank then

$$0 = \mathbf{A}\underline{x} = \sum_i x_i \underline{a}_{*i} \ \Leftrightarrow \ \underline{x} = \underline{0}.$$

If in addition $\mathbf{A}$ is square then it is said to be non-singular.

2.3.4 MATRIX INVERSION

If $\mathbf{A}$ is non-singular square matrix then it has an inverse $\mathbf{A}^{-1}$ which satisfies the relation $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. In the special case of a 2×2 matrix the matrix inverse is given by (Cramèr's formula)

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad \text{if} \quad ad \neq bc$$

Sometimes when a matrix has special structure its inverse has a simple form. The books by Graybill [21] and Golub and VanLoan [19] give many interesting and useful examples. Some results which we will need in this text are: the *Sherman-Morrison-Woodbury identity*

$$[\mathbf{A} + \mathbf{U}\mathbf{V}^T]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}[\mathbf{I} + \mathbf{V}^T\mathbf{A}^{-1}\mathbf{U}]^{-1}\mathbf{V}^T\mathbf{A}^{-1}, \tag{1}$$

where $\mathbf{A}, \mathbf{U}, \mathbf{V}$ are compatible matrices, $[\mathbf{A} + \mathbf{U}\mathbf{V}^T]^{-1}$ and $\mathbf{A}^{-1}$ exist; and the *partitioned matrix inverse identity*

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} [\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}]^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}[\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}]^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}[\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}]^{-1} & [\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}]^{-1} \end{bmatrix}, \tag{2}$$

assuming that all the indicated inverses exist.

2.3.5 ORTHOGONAL AND UNITARY MATRICES

A real square matrix $\mathbf{A}$ is said to be orthogonal if all of its columns are orthonormal, i.e.,

$$\mathbf{A}^T\mathbf{A} = \mathbf{I}. \tag{3}$$

The generalization of orthogonality to complex matrices $\mathbf{A}$ is the property of being unitary,

$$\mathbf{A}^H\mathbf{A} = \mathbf{I}.$$

The relation (3) implies that if $\mathbf{A}$ is an orthogonal matrix it is invertible and has a very simple inverse

$$\mathbf{A}^{-1} = \mathbf{A}^T.$$

2.3.6 GRAMM-SCHMIDT ORTHOGONALIZATION AND ORTHONORMALIZATION

Let $\underline{x}_1, \ldots, \underline{x}_n$ be a set of n linearly independent p dimensional column vectors ($n \leq p$) whose linear span is the subspace $\mathcal{H}$. Gramm-Schmidt orthogonalization is an algorithm that can be applied to this set of vectors to obtain a set of n orthogonal vectors $\underline{y}_1, \ldots, \underline{y}_n$ that spans the same subspace. This algorithm proceeds as follows.

Step 1: select $\underline{y}_1$ as an arbitrary starting point in $\mathcal{H}$. For example, choose any coefficient vector $\underline{a}_1 = [a_{11}, \ldots, a_{1n}]^T$ and define $\underline{y}_1 = \mathbf{X}\underline{a}_1$ where $\mathbf{X} = [\underline{x}_1, \ldots, \underline{x}_n]$.

Step 2: construct the other $n-1$ vectors $\underline{y}_2, \ldots, \underline{y}_n$ by the following recursive procedure:

`For` $j = 2, \ldots, n$: $\underline{y}_j = \underline{x}_j - \sum_{i=1}^{j} K_i \underline{y}_{i-1}$ `where` $K_j = \underline{x}_j^T \underline{y}_{j-1} / \underline{y}_{j-1}^T \underline{y}_{j-1}$.

The above Gramm-Schmidt procedure can be expressed in compact matrix form [60]

$$\mathbf{Y} = \mathbf{H}\mathbf{X},$$

where $\mathbf{Y} = [\underline{y}_1, \ldots, \underline{y}_n]$ and $\mathbf{H}$ is called the Gramm-Schmidt matrix.

If after each step $j = 1, \ldots, n$ of the procedure one maps normalizes the length of $\underline{y}_j$, i.e., $\underline{y}_j \leftarrow \tilde{\underline{y}}_j = \underline{y}_j / \|\underline{y}_j\|$, the algorithm produces an orthonormal set of vectors. This is called Gram-Schmidt orthonormalization and produces an matrix $\tilde{\mathbf{Y}}$ with orthonormal columns and identical column span as that of $\mathbf{X}$. The Gramm-Schmidt orthonormalization procedure is often used to generate an orthonormal basis $\underline{y}_1, \ldots, \underline{y}_p]$ for $\mathbb{R}^p$ starting from an arbitrarily selected initial vector $\underline{y}_1$. The matrix formed from such a basis will have the structure

$$\mathbf{Y} = \begin{bmatrix} \underline{y}_1 \\ \underline{v}_2 \\ \vdots \\ \underline{v}_n \end{bmatrix}$$

and

$$\mathbf{Y}^T \mathbf{Y} = \mathbf{I}.$$

In the above $\underline{v}_2, \ldots, \underline{v}_n$ are orthonormal vectors that are said to accomplish *completion of the basis* with respect to the initial vector $\underline{y}_1$.

2.3.7 EIGENVALUES OF A SYMMETRIC MATRIX

If $\mathbf{R}$ is arbitrary $n \times n$ **symmetric matrix**, that is, $\mathbf{R}^T = \mathbf{R}$, then there exist a set of n orthonormal eigenvectors $\underline{\nu}_i$,

$$\underline{\nu}_i^T \underline{\nu}_j = \Delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

and a set of associated eigenvectors λ_i such that:

$$\mathbf{R}\underline{\nu}_i = \lambda_i \underline{\nu}_i, \qquad i = 1, \ldots, n.$$

These eigenvalues and eigenvectors satisfy:

$$\begin{aligned} \underline{\nu}_i^T \mathbf{R} \underline{\nu}_i &= \lambda_i \\ \underline{\nu}_i^T \mathbf{R} \underline{\nu}_j &= 0, \quad i \neq j. \end{aligned}$$

2.3.8 MATRIX DIAGONALIZATION AND EIGENDECOMPOSITION

Let $\mathbf{U} = [\underline{\nu}_1, \ldots, \underline{\nu}_n]$ be the $n \times n$ matrix formed from the eigenvectors of a symmetric matrix $\mathbf{R}$. If $\mathbf{R}$ is real symmetric $\mathbf{U}$ is a real orthogonal matrix while if $\mathbf{R}$ is complex Hermitian symmetric $\mathbf{U}$ is a complex unitary matrix:

$$\begin{aligned} \mathbf{U}^T \mathbf{U} &= \mathbf{I}, \quad (\mathbf{U} \text{ an orthogonal matrix}) \\ \mathbf{U}^H \mathbf{U} &= \mathbf{I}, \quad (\mathbf{U} \text{ a unitary matrix}). \end{aligned}$$

where as before H denotes Hermitian transpose. As the Hermitian transpose of a real matrix is equal to its ordinary transpose, we will use the more general notation $\mathbf{A}^H$ for any (real or complex) matrix $\mathbf{A}$.

The matrix $\mathbf{U}$ can be used to diagonalize $\mathbf{R}$

$$\mathbf{U}^H \mathbf{R} \mathbf{U} = \mathbf{\Lambda}, \tag{4}$$

In cases of both real and Hermitian symmetric $\mathbf{R}$ the matrix $\mathbf{\Lambda}$ is diagonal and real valued

$$\mathbf{\Lambda} = \mathrm{diag}(\lambda_i) = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix},$$

where λ_i's are the eigenvalues of $\mathbf{R}$.

The expression (4) implies that

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H,$$

which is called the *eigendecomposition* of $\mathbf{R}$. As $\mathbf{\Lambda}$ is diagonal, an equivalent summation form for this eigendecomposition is

$$\mathbf{R} = \sum_{i=1}^{n} \lambda_i \underline{\nu}_i \underline{\nu}_i^H. \tag{5}$$

2.3.9 QUADRATIC FORMS AND NON-NEGATIVE DEFINITE MATRICES

For a square symmetric matrix $\mathbf{R}$ and a compatible vector $\underline{x}$, a quadratic form is the scalar defined by $\underline{x}^T\mathbf{R}\underline{x}$. The matrix $\mathbf{R}$ is non-negative definite (nnd) if for any $\underline{x}$

$$\underline{x}^T\mathbf{R}\underline{x} \geq 0. \tag{6}$$

$\mathbf{R}$ is positive definite (pd) if it is nnd and "=" in (6) implies that $\underline{x} = \underline{0}$, or more explicitly $\mathbf{R}$ is pd if

$$\underline{x}^T\mathbf{R}\underline{x} > 0, \quad \underline{x} \neq \underline{0}. \tag{7}$$

Examples of nnd (pd) matrices:

* $\mathbf{R} = \mathbf{B}^T\mathbf{B}$ for arbitrary (pd) matrix $\mathbf{B}$

* $\mathbf{R}$ symmetric with only non-negative (positive) eigenvalues

Rayleigh Theorem: If $\mathbf{A}$ is a nnd $n \times n$ matrix with eigenvalues $\{\lambda_i\}_{i=1}^n$ the quadratic form

$$\min(\lambda_i) \leq \frac{\underline{u}^T\mathbf{A}\underline{u}}{\underline{u}^T\underline{u}} \leq \max(\lambda_i)$$

where the lower bound is attained when $\underline{u}$ is the eigenvector of $\mathbf{A}$ associated with the minimum eigenvalue of $\mathbf{A}$ and the upper bound is attained by the eigenvector associated with the maximum eigenvalue of $\mathbf{A}$.

2.4 POSITIVE DEFINITENESS OF SYMMETRIC PARTITIONED MATRICES

If $\mathbf{A}$ is a symmetric matrix with partition representation (2) then it is easily shown that

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{O}^T \\ \mathbf{O} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O}^T \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix}^{-1}, \tag{8}$$

as long as $\mathbf{A}_{22}^{-1}$ exists. Here $\mathbf{O}$ denotes a block of zeros. This implies: if $\mathbf{A}$ is positive definite the matrices $\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$ and $\mathbf{A}_{22}$ are pd. By using an analogous identity we can conclude that $\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$ and $\mathbf{A}_{11}$ are also pd.

2.4.1 DETERMINANT OF A MATRIX

If $\mathbf{A}$ is any square matrix its determinant is

$$|\mathbf{A}| = \prod_i \lambda_i$$

Note: a square matrix is non-singular iff its determinint is non-zero.
If $\mathbf{A}$ is partitioned as in (2) and $\mathbf{A}_{11}^{-1}$ and $\mathbf{A}_{22}^{-1}$ exist then

$$|\mathbf{A}| = |\mathbf{A}_{11}||\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}| = |\mathbf{A}_{22}||\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}| \tag{9}$$

This follows from the decomposition (8).

2.4.2 TRACE OF A MATRIX

For any square matrix $\mathbf{A} = ((a_{ij}))$ the trace of $\mathbf{A}$ is defined as

$$\text{trace}\{\mathbf{A}\} = \sum_i a_{ii} = \sum_i \lambda_i$$

One has an important identity: for compatible matrices $\mathbf{A}$ and $\mathbf{B}$

$$\text{trace}\{\mathbf{AB}\} = \text{trace}\{\mathbf{BA}\}.$$

This has the following implication for quadratic forms:

$$\underline{x}^T\mathbf{R}\underline{x} = \text{trace}\{\underline{x}\underline{x}^T\ \mathbf{R}\}.$$

2.4.3 VECTOR DIFFERENTIATION

Differentiation of functions of a vector variable often arise in signal processing and estimation theory. If $\underline{h} = [h_1, \ldots, h_n]^T$ is an $n \times 1$ vector and $g(\underline{h})$ is a scalar function then the gradient of $g(\underline{h})$, denoted $\nabla g(\underline{h})$ or $\nabla_{\underline{h}} g(\underline{h})$ when necessary for conciseness, is defined as the (column) vector of partials

$$\nabla g = \left[\frac{\partial g}{\partial h_1}, \ldots, \frac{\partial g}{\partial h_n}\right]^T.$$

In particular, if c is a constant

$$\nabla_{\underline{h}} c = \underline{0},$$

if $\underline{x} = [x_1, \ldots, x_n]^T$

$$\nabla_{\underline{h}}(\underline{h}^T\underline{x}) = \nabla_{\underline{h}}(\underline{x}^T\underline{h}) = \underline{x},$$

and if $\mathbf{B}$ is an $n \times n$ matrix

$$\nabla_{\underline{h}}(\underline{h} - \underline{x})^T\mathbf{B}(\underline{h} - \underline{x}) = 2\mathbf{B}(\underline{h} - \underline{x}).$$

For a vector valued function $\underline{g}(\underline{h}) = [g_1(\underline{h}), \ldots, g_m(\underline{h})]^T$ the gradient of $\underline{g}(\underline{h})$ is an $m \times n$ matrix. In particular, for a scalar function $g(\underline{h})$, the two applications of the gradient $\nabla(\nabla g)^T$ gives the $n \times n$ Hessian matrix of g, denoted as $\nabla^2 g$. This yields useful and natural identities such as:

$$\nabla_{\underline{h}}^2(\underline{h} - \underline{x})^T\mathbf{B}(\underline{h} - \underline{x}) = 2\mathbf{B}.$$

For a more detailed discussion of vector differentiation the reader is referred to Kay [36].

2.5 SIGNALS AND SYSTEMS BACKGROUND

Here we review some of the principal results that will be useful for dealing with signals and systems encountered in this book.

2.5.1 GEOMETRIC SERIES

One of the most useful formulas in discrete time signal and systems engineering is:

$$\sum_{i=0}^{n} a^n = \frac{1-a^{n+1}}{1-a}, \quad \text{if } a \neq 1; \qquad \sum_{i=0}^{\infty} a^n = \frac{1}{1-a}, \quad \text{if } |a| < 1.$$

2.5.2 LAPLACE AND FOURIER TRANSFORMS OF FUNCTIONS OF A CONTINUOUS VARIABLE

If $h(t)$, $-\infty < t < \infty$, a square integrable function of a continuous variable t (usually time) then its Laplace and Fourier transforms are defined as follows.

The Laplace transform of h is

$$\mathcal{L}\{h\} = H(s) = \int_{-\infty}^{\infty} h(t) e^{-st} \, dt$$

where $s = \sigma + j\omega \in \mathbb{C}$ is a complex variable.

The Fourier transform of h is

$$\mathcal{F}\{h\} = H(\omega) = \int_{-\infty}^{\infty} h(t) e^{-j\omega t} \, dt$$

Note: $\mathcal{F}\{h\} = \mathcal{L}\{h\}|_{s=j\omega}$.

Example: if $h(t) = e^{-at}u(t)$, for $a > 0$, then the Laplace transform is

$$H(s) = \int_0^{\infty} e^{-at} e^{-st} \, dt = \int_0^{\infty} e^{-(a+s)t} \, dt = \frac{-1}{a+s} e^{-(a+st)} \Big|_0^{\infty} = \frac{1}{a+s}$$

2.5.3 Z-TRANSFORM AND DISCRETE-TIME FOURIER TRANSFORM (DTFT)

If h_k, $k = \ldots, -1, 0, 1, \ldots$, is a square summable function of a discrete variable then its Z-transform and discrete-time Fourier transform (DTFT) are defined as follows.

The Z-transform is

$$\mathcal{Z}\{h\} = H(z) = \sum_{k=-\infty}^{\infty} h_k z^{-k}$$

The DTFT is

$$\mathcal{F}\{h\} = H(\omega) = \sum_{k=-\infty}^{\infty} h_k e^{-j\omega k}$$

Note: $H(\omega)$ really means $H(e^{j\omega})$ and is an abuse of notation

- $\mathcal{F}\{h\} = \mathcal{Z}\{h\}|_{z=e^{j\omega}}$
- the DTFT is always periodic in ω with period 2π.

Example: if $h_k = a^{|k|}$, then for $|az^{-1}| < 1$ and $|az| < 1$, the Z-transform is

$$H(z) = \sum_{k=-\infty}^{\infty} a^{|k|} z^{-k} = \sum_{k=-\infty}^{-1} a^{-k} z^{-k} + \sum_{k=0}^{\infty} a^k z^{-k}$$

$$= \sum_{k=1}^{\infty} (az)^k + \sum_{k=0}^{\infty} (az^{-1})^k = \frac{az}{1-az} + \frac{1}{1-az^{-1}}$$

Likewise the DTFT is (for $|a| < 1$):

$$H(\omega) = H(z)|_{z=e^{j\omega}} = \frac{1-a^2}{1-2a\cos\omega + a^2}$$

2.5.4 CONVOLUTION: CONTINUOUS TIME

If $h(t)$ and $x(t)$ are square integrable functions of a continuous variable t then the convolution of x and h is defined as

$$(h * x)(t) = \int_{-\infty}^{\infty} h(t-\tau)x(\tau)\, d\tau$$

Note: The convolution of h and x is a waveform indexed by time t. $(h*x)(t)$ is this waveform evaluated at time t and is frequently denoted $h(t) * x(t)$.

Example: $h(t) = e^{-at}u(t)$, for $a > 0$, (the filter) and $x(t) = e^{-bt}u(t)$, for $b > 0$, (the filter input) then

$$(h * x)(t) = \int_{-\infty}^{\infty} e^{-a(t-\tau)} e^{-b\tau} u(t-\tau)u(\tau)\, d\tau = \left(\int_0^t e^{-a(t-\tau)} e^{-b\tau}\, d\tau\right) u(t)$$

$$= e^{-at}\left(\int_0^t e^{-(b-a)\tau}\, d\tau\right) u(t) = e^{-at}\left(\frac{-1}{b-a} e^{-(b-a)\tau}\Big|_0^t\right) u(t) = \frac{e^{-at} - e^{-bt}}{b-a} u(t)$$

2.5.5 CONVOLUTION: DISCRETE TIME

If h_k and x_k are square integrable sequences then

$$h_n * x_n = \sum_{j=-\infty}^{\infty} h_j x_{n-j} = \sum_{j=-\infty}^{\infty} h_{n-j} x_j$$

h_k is a called a "causal" filter if it is zero for negative indices:

$$h_k = 0, \quad k < 0$$

2.5.6 CORRELATION: DISCRETE TIME

For time sequences $\{x_k\}_{k=1}^n$ and $\{y_k\}_{k=1}^n$ their temporal correlation is

$$z_n = \sum_{j=1}^n x_k y_k^*$$

2.5.7 RELATION BETWEEN CORRELATION AND CONVOLUTION

$$z_n = \sum_{j=1}^n x_k y_k^* = \sum_{j=-\infty}^{\infty} x_k h_{n-k} = h_n \star x_n$$

where

$$h_k = \begin{cases} y_{n-k}^*, & k = 1, \ldots, n \\ 0, & o.w. \end{cases}$$

2.5.8 CONVOLUTION AS A MATRIX OPERATION

Let h_k be a causal filter and let x_k be an input starting at time $k = 1$. Arranging n outputs z_k in a vector $\underline{z}$ it is easy to see that

$$\begin{aligned} \underline{z} = \begin{bmatrix} z_n \\ \vdots \\ z_1 \end{bmatrix} &= \begin{bmatrix} \sum_{j=1}^n h_{n-j} x_j \\ \vdots \\ \sum_{j=1}^n h_{1-j} x_j \end{bmatrix} \\ &= \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ 0 & h_0 & \ddots & h_{n-2} \\ \vdots & \ddots & h_0 & h_1 \\ 0 & \cdots & 0 & h_0 \end{bmatrix} \begin{bmatrix} x_n \\ \vdots \\ x_1 \end{bmatrix} \end{aligned}$$

2.6 BACKGROUND REFERENCES

There are many useful textbooks that cover areas of this chapter. I learned elementary linear algebra from Noble and Daniel [52]. A more advanced book that is focused on computational linear algebra is Golub and Van Loan [19] which covers many fast and numerically stable algorithms arising in signal processing. Another nice book on linear algebra with emphasis on statistical applications is Graybill [21] that contains lots of useful identities for multivariate Gaussian models. For background on signals and systems Oppenheim and Willsky [53] and Proakis and Manolakis [56] are good elementary textbooks. The encyclopedic book by Moon and Stirling [49] is a good general resource for mathematical methods in signal processing.

2.7 EXERCISES

2.1 Let $\underline{a}$, $\underline{b}$ be $n \times 1$ vectors and let $\mathbf{C}$ be an invertible $n \times n$ matrix. Assuming α is not equal to $-1/(\underline{a}^T \mathbf{C}^{-1} \underline{b})$ show the following identity

$$[\mathbf{C} + \alpha \underline{a}\underline{b}^T]^{-1} = \mathbf{C}^{-1} - \mathbf{C}^{-1}\underline{a}\underline{b}^T \mathbf{C}^{-1} \alpha/(1 + \alpha \underline{a}^T \mathbf{C}^{-1} \underline{b}).$$

2.2 A discrete time LTI filter $h(k)$ is causal when $h(k) = 0,\ k < 0$ and anticausal when $h(k) = 0,\ k > 0$. Show that if $|h(k)| < \infty$ for all k, the transfer function $H(z) = \sum_{k=-\infty}^{\infty} h(k) z^{-k}$ of a causal LTI has no singularities outside the unit circle, i.e. $|H(z)| < \infty,\ |z| > 1$ while an anticausal LTI has no singularities inside the unit circle, i.e. $|H(z)| < \infty,\ |z| < 1$. (Hint: generalized triangle inequality $|\sum_i a_i| \leq \sum |a_i|$)

2.3 A discrete time LTI filter $h(k)$ is said to be BIBO stable when $\sum_{k=-\infty}^{\infty} |h(k)| < \infty$. Define the transfer function (Z-transform) $H(z) = \sum_{k=-\infty}^{\infty} h(k) z^{-k}$, for z a complex variable.

(a) Show that $H(z)$ has no singularities on the unit circle, i.e $|H(z)| < \infty,\ |z| = 1$.

(b) Show that if a BIBO stable $h(k)$ is causal then $H(z)$ has all its singularities (poles) strictly inside the unit circle, i.e $|H(z)| < \infty,\ |z| \geq 1$.

(c) Show that if a BIBO stable $h(k)$ is anticausal, i.e. $h(k) = 0,\ k > 0$, then $H(z)$ has all its singularities (poles) strictly outside the unit circle, i.e $|H(z)| < \infty,\ |z| \leq 1$.

2.4 If you are only given the mathematical form of the transfer function $H(z)$ of an LTI, and not told whether it corresponds to an LTI which is causal, anticausal, or stable, then it is not possible to uniquely specify the impulse response $\{h_k\}_k$. This simple example illustration this fact. The regions $\{z : |z| > a\}$ and $\{z : |z| \leq a\}$, specified in (a) and (b) are called the regions of convergence of the filter and specify whether the filter is stable, causal or anticausal.

Let $H(z)$ be

$$H(z) = \frac{1}{1 - az^{-1}}$$

(a) Show that if the LTI is causal, then for $|z| > |a|$ you can write $H(z)$ as the convergent series

$$H(z) = \sum_{k=0}^{\infty} a^k z^{-k}, \quad |z| > |a|$$

which corresponds to $h_k = a^k,\ k = 0, 1, \ldots$ and $h_k = 0,\ k < 0$.

(b) Show that if the LTI is anticausal, then for $|z| < |a|$ you can write $H(z)$ as the convergent series

$$H(z) = -\sum_{k=0}^{\infty} a^{-k} z^{k+1}, \quad |z| < |a|$$

which corresponds to $h_k = -a^{-k},\ k = 1, 2 \ldots$ and $h_k = 0,\ k \geq 0$.

(c) Show that if $|a| < 1$ then the causal LTI is BIBO stable while the anti-causal LTI is BIBO unstable while if $|a| > 1$ then the reverse is true. What happens to stability when $|a| = 1$?

2.5 An LTI has transfer function

$$H(z) = \frac{3 - 4z^{-1}}{1 - 3.5z^{-1} + 1.5z^{-2}}$$

(a) If you are told that the LTI is stable specify the region of convergence (ROC) in the z-plane, i.e. specify the range of values of $|z|$ for which $|H(z)| < \infty$, and specify the impulse response.

(b) If you are told that the LTI is causal specify the region of convergence (ROC) in the z-plane, and specify the impulse response.

(c) If you are told that the LTI is anticausal specify the region of convergence (ROC) in the z-plane, and specify the impulse response.

End of chapter

3 STATISTICAL MODELS

Keywords: sampling distributions, sufficient statistics, exponential families.

Estimation, detection and classification can be grouped under the broad heading of statistical inference which is the process of inferring properties about the distribution of a random variable X given a realization x, which is also called a data sample, a measurement, or an observation. A key concept is that of the statistical model which is simply a hypothesized probability distribution or density function $f(x)$ for X. Broadly stated statistical inference explores the possibility of fitting a given model to the data x. To simplify this task it is common to restrict $f(x)$ to a class of *parameteric models* $\{f(x;\underline{\theta})\}_{\underline{\theta}\in\Theta}$, where $f(x;\bullet)$ is a known function and $\underline{\theta}$ is a vector of unknown parameters taking values in a parameter space Θ. In this special case statistical inference boils down to inferring properties of the *true value* of $\underline{\theta}$ parameterizing $f(x;\underline{\theta})$ that generated the data sample x.

In this chapter we discuss several models that are related to the ubiquitous Gaussian distribution, the more general class of exponential families of distributions, and the important concept of a sufficient statistic for infering properties about $\underline{\theta}$.

3.1 THE GAUSSIAN DISTRIBUTION AND ITS RELATIVES

The Gaussian distribution and its close relatives play a major role in parameteric statistical inference due to the relative simplicity of the Gaussian model and its broad applicability (recall the Central Limit Theorem!). Indeed, in engineering and science the Gaussian distribution is probably the most commonly invoked distribution for random measurements. The Gaussian distribution is also called the Normal distribution. The probability density function (pdf) of a Gaussian random variable (rv) X is parameterized by two parameters, θ_1 and θ_2, which are the location parameter, denoted μ ($\mu \in \mathbb{R}$), and the (squared) scale parameter, denoted σ^2 ($\sigma^2 > 0$). The pdf of this Gaussian rv has the form

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

When $\mu = 0$ and $\sigma^2 = 1$, X is said to be a standard Gaussian (Normal) rv. A Gaussian random variable with location parameter μ and scale parameter $\sigma > 0$ can be represented by

$$X = \sigma Z + \mu, \tag{10}$$

where Z is a standard Gaussian rv.

The cumulative density function (cdf) of a standard Gaussian random variable Z is denoted $\mathcal{N}(z)$ and is defined in the conventional manner

$$\mathcal{N}(z) = P(Z \le z).$$

Equivalently,

$$\mathcal{N}(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv.$$

Using (10) the cdf of a non-standard Gaussian rv X with parameters μ and σ^2 can be expressed in terms of the cdf $\mathcal{N}(z)$ of a standard Gaussian rv Z:

$$P(X \le x) = P(\underbrace{(X-\mu)/\sigma}_{Z} \le (x-\mu)/\sigma) = \mathcal{N}\left(\frac{x-\mu}{\sigma}\right)$$

The standard Normal cdf $\mathcal{N}(x)$ can be related to the error function or error integral [1]: $\mathrm{erf}(u) = \frac{2}{\sqrt{\pi}}\int_0^u e^{-t^2}dt$, $x \geq 0$, through the relation

$$\mathcal{N}(x) = \begin{cases} \frac{1}{2}[1+\mathrm{erf}(|x|/\sqrt{2})] & x \geq 0 \\ \frac{1}{2}[1-\mathrm{erf}(|x|/\sqrt{2})], & x < 0 \end{cases}.$$

For positive integer order ν, the moments of a standard Gaussian random variable Z are [30, 13.3]

$$E[Z^\nu] = \begin{cases} (\nu-1)(\nu-3)\cdots 3\cdot 1, & \nu \text{ even} \\ 0, & \nu \text{ odd} \end{cases}$$

where $E[g(Z)] = \int_{-\infty}^{\infty} g(z)f(z)dz$ denotes statistical expectation of the rv $g(Z)$ under the pdf $f(z)$ for rv Z. These moment relations can easily be derived by looking at the coefficients of $(ju)^k/k!$, $k = 1, 2, \ldots$ in the power series expansion about $ju = 0$ of the characteristic function $\Phi_Z(u) = E[e^{juZ}] = e^{-u^2/2}$.

In particular, using (10), this implies that the first and second moments of a non-standard Gaussian rv X are $E[X] = \mu$ and $E[X^2] = \mu^2 + \sigma^2$, respectively. Thus for a Gaussian rv X we can identify the (ensemble) mean $E[X] = \mu$ and variance $\mathrm{var}(X) = E[(X - E[X])^2] = E[X^2] - E^2[X] = \sigma^2$ as the location and (squared) scale parameters, respectively, of the pdf $f(x;\mu,\sigma^2)$ of X. In the sequel we will need the following expression for the (non-central) mean deviation $E[|X + a|]$ for Gaussian X [31, 29.6]:

$$E[|X+a|] = \sqrt{\frac{2}{\pi}}e^{-a^2/2} + a(1 - 2\mathcal{N}(-a)). \tag{11}$$

In referring to rv's and operations on rv's in this book the following compact notations are sometimes used:

* "X is distributed as a Gaussian random variable with mean μ and variance σ^2"

$$X \sim \mathcal{N}(\mu, \sigma^2) \tag{12}$$

* "X is equal to a scaled and shifted standard Gaussian random variable"

$$X = a \underbrace{Z}_{\mathcal{N}(0,1)} + b \Leftrightarrow X \sim \mathcal{N}(b, a^2)$$

or, in shorthand notation,

$$X = a\,\mathcal{N}(0,1) + b \Leftrightarrow X \sim \mathcal{N}(b, a^2). \tag{13}$$

For example, in the following shorthand notation $X_1, \ldots, X_n$ are independent identically distributed (iid) $\mathcal{N}(0,1)$ rv's

$$\sum_{i=1}^n \mathcal{N}(0,1) = \sum_{i=1}^n X_i.$$

Note that the above is an abuse of notation since $\mathcal{N}(0,1)$ is being used to denote both a Gaussian probability distribution in (12) and a Gaussian random variable in (13). As in all abuses of this type the ambiguity is resolved from the context: we will never write $\mathcal{N}(0,1)$ into an algebraic or other type of equation like the one in (13) when $\mathcal{N}(0,1)$ is meant to denote a Gaussian distribution function as opposed to a Gaussian random variable.

Other notational shortcuts are the following. When we write

$$\mathcal{N}(v) = \alpha$$

we mean that "the cdf of a $\mathcal{N}(0,1)$ rv equals α when evaluated at a point $v \in \mathbb{R}$." Likewise

$$\mathcal{N}^{-1}(\alpha) = v$$

is to be read as "the inverse cdf of a $\mathcal{N}(0,1)$ rv equals v when evaluated at a point $\alpha \in [0,1]$." Finally, by

$$\underline{X} \sim \mathcal{N}_n(\underline{\mu},\, \mathbf{R})$$

we mean "$\underline{X}$ is distributed as an n-dimensional Gaussian random vector with mean $\underline{\mu}$ and covariance matrix $\mathbf{R}$"

3.1.1 MULTIVARIATE GAUSSIAN DISTRIBUTION

When one passes an i.i.d. Gaussian random sequence through a linear filter the output remains Gaussian but is no longer i.i.d; the filter smooths the input and introduces correlation. Remarkably, if the input to the filter is Gaussian then the output is also Gaussian, i.e., the joint distribution of any p samples of the output is multivariate Gaussian. To be specific, a random vector $\underline{X} = [X_1, \ldots, X_p]^T$ is multivariate Gaussian with mean parameter $\underline{\mu}$ and covariance matrix parameter $\mathbf{\Lambda}$ if it has a joint density of the form

$$f(\underline{x}) = \frac{1}{(2\pi)^{p/2}|\mathbf{\Lambda}|^{1/2}} \exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu})\mathbf{\Lambda}^{-1}(\underline{x}-\underline{\mu})\right) \quad \underline{x} \in \mathbb{R}^p. \tag{14}$$

where $|\mathbf{\Lambda}|$ denotes the the determinant of $\mathbf{\Lambda}$. The p-variate Gaussian distribution depends on $p(p+3)/2$ parameters, which we can concatenate into a parameter vector $\underline{\theta}$ consisting of the p elements of the mean vector

$$\underline{\mu} = [\mu_1, \ldots, \mu_p]^T = E[\underline{X}],$$

and the $p(p+1)/2$ distinct parameters of the symmetric positive definite $p \times p$ covariance matrix

$$\mathbf{\Lambda} = \mathrm{cov}(\underline{Z}) = E\left[(\underline{Z}-\underline{\mu})(\underline{Z}-\underline{\mu})^T\right].$$

Some useful facts about the multivariate Gaussian random variables are (for derivations of these properties see Morrison [50]):

• **Unimodality and symmetry of the Gaussian density**: The multivariate Gaussian density (14) is unimodal (has a unique maximum) and is symmetric about its mean parameter.

• **Uncorrelated Gaussians are independent**: When the covariance matrix $\mathbf{\Lambda}$ is diagonal, i.e., $\mathrm{cov}(X_i, X_j) = 0$, $i \neq j$, then the multivariate Gaussian density reduces to a product of univariate densities

$$f(\underline{X}) = \prod_{i=1}^{n} f(X_i)$$

where

$$f(X_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{1}{2\sigma_i^2}(X_i-\mu_i)^2}$$

is the univariate Gaussian density with $\sigma_i^2 = \text{var}(X_i)$. Thus uncorrelated Gaussian random variables are in fact independent random variables.

• **Marginals of a Gaussian density are Gaussian**: If $\underline{X} = [X_1, \ldots, X_m]^T$ is multivariate Gaussian then any subset of the elements of $\underline{X}$ is also Gaussian. In particular X_1 is univariate Gaussian and $[X_1, X_2]$ is bivariate Gaussian.

• **Linear combination of Gaussian random variables are Gaussian**: Let $\underline{X} = [X_1, \ldots, X_m]^T$ be a multivariate Gaussian random vector and let $\mathbf{H}$ be a $p \times m$ non-random matrix. Then $\underline{Y} = \mathbf{H}\underline{X}$ is a vector of linear combinations of the X_i's. The distribution of $\underline{Y}$ is multivariate (p-variate) Gaussian with mean $\underline{\mu}_Y = E[\underline{Y}] = \mathbf{H}\underline{\mu}$ and $p \times p$ covariance matrix $\mathbf{\Lambda}_Y = \text{cov}(\underline{Y}) = \mathbf{H}\text{cov}(\underline{X})\mathbf{H}^T$.

• **A vector of i.i.d. zero mean Gaussian random variables is invariant to rotation**: Let $\underline{X} = [X_1, \ldots, X_m]^T$ be vector of zero mean Gaussian random variables with covariance $\text{cov}(\underline{X}) = \sigma^2\mathbf{I}$. If $\mathbf{U}$ is an orthogonal $m \times m$ matrix, i.e., $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, then $\underline{Y} = \mathbf{U}^T\underline{X}$ has the same distribution as $\underline{X}$.

• **The conditional distribution of a Gaussian given another Gaussian is Gaussian**: Let the vector $\underline{Z}^T = [\underline{X}^T, \underline{Y}^T] = [X_1, \ldots, X_p, Y_1, \ldots, Y_q]^T$ be multivariate ($(p+q)$-variate) Gaussian with mean parameters $\underline{\mu}_Z^T = [\underline{\mu}_X^T, \underline{\mu}_Y^T]$ and covariance parameters $\mathbf{\Lambda}_Z$. Then the conditional density $f_{Y|X}(\underline{y}|\underline{x})$ of $\underline{Y}$ given $\underline{X} = \underline{x}$ is multivariate (q-variate) Gaussian of the form (14) with mean and covariance parameters $\underline{\mu}$ and $\mathbf{\Lambda}$ respectively given by (15) and (16) below.

• **Conditional mean of a Gaussian given another Gaussian is linear and conditional covariance is constant**: For the aforementioned multivariate Gaussian vector $\underline{Z}^T = [\underline{X}^T, \underline{Y}]^T$ partition its covariance matrix as follows

$$\mathbf{\Lambda}_Z = \begin{bmatrix} \mathbf{\Lambda}_X & \mathbf{\Lambda}_{X,Y} \\ \mathbf{\Lambda}_{X,Y}^T & \mathbf{\Lambda}_Y \end{bmatrix},$$

where $\mathbf{\Lambda}_X = \text{cov}(\underline{X}) = E[(\underline{X} - \underline{\mu}_X)(\underline{X} - \underline{\mu}_X)^T]$ is $p \times p$, $\mathbf{\Lambda}_Y = \text{cov}(\underline{Y}) = E[(\underline{Y}\underline{\mu}_Y)(\underline{Y} - \underline{\mu}_Y)^T]$ is $q \times q$, and $\mathbf{\Lambda}_{X,Y} = \text{cov}_{\underline{\theta}}(\underline{X}, \underline{Y}) = E[(\underline{X} - \underline{\mu}_X)(\underline{Y} - \underline{\mu}_Y)^T]$ is $p \times q$. The mean of the multivariate Gaussian conditional density $f(\underline{y}|\underline{x})$, the conditional mean, is linear in $\underline{x}$

$$\underline{\mu}_{Y|X}(\underline{x}) = E[\underline{Y}|\underline{X} = \underline{x}] = \underline{\mu}_Y + \mathbf{\Lambda}_{X,Y}^T \mathbf{\Lambda}_X^{-1}(\underline{x} - \underline{\mu}_X) \tag{15}$$

and the conditional covariance does not depend on $\underline{x}$

$$\mathbf{\Lambda}_{Y|X} = \text{cov}(\underline{Y}|\underline{X} = x) = \mathbf{\Lambda}_Y - \mathbf{\Lambda}_{X,Y}^T \mathbf{\Lambda}_X^{-1} \mathbf{\Lambda}_{X,Y}. \tag{16}$$

3.1.2 CENTRAL LIMIT THEOREM

One of the most useful results in statistics is the central limit theorem, abbreviated to CLT. This theorem allows one to approximate the distribution of sums of i.i.d. finite variance random variables by a Gaussian distribution. Below we give a general version of the CLT that applies to vector valued r.v.s. For a simple proof of the scalar case see Mood, Graybill and Boes [48]. For proof in the multivariate case see Serfling [Ch. 1][62], which also covers the CLT for the non i.i.d. case.

(Lindeberg-Lévy) Central Limit Theorem: Let $\{\underline{X}_i\}_{i=1}^n$ be i.i.d. random vectors in $\mathbb{R}^p$ with common mean $E[\underline{X}_i] = \underline{\mu}$ and finite positive definite covariance matrix $\text{cov}(\underline{X}_i) = \mathbf{\Lambda}$. Then as n goes to infinity the distribution of the random vector $\underline{Z}_n = n^{-1/2} \sum_{i=1}^n (\underline{X}_i - \underline{\mu})$ converges to a p-variate Gaussian distribution with zero mean and covariance $\mathbf{\Lambda}$.

The CLT can also be expressed in terms of the sample mean $\overline{\underline{X}} = \overline{\underline{X}}(n) = n^{-1} \sum_{i=1}^n \underline{X}_i$: as $n \to \infty$

$$\sqrt{n}(\overline{\underline{X}}(n) - \underline{\mu}) \longrightarrow \underline{Z}$$

where $\underline{Z}$ is a zero mean Gaussian random vector with covariance matrix $\mathbf{\Lambda}$. Thus, for large but finite n, $\overline{\underline{X}}$ is approximately Gaussian

$$\overline{\underline{X}} \approx (\underline{Z}/\sqrt{n} + \underline{\mu}),$$

with mean $\underline{\mu}$ and covariance $\mathbf{\Lambda}/n$. For example, in the case of a scalar X_i, the CLT gives the useful large n approximation

$$P(n^{-1} \sum_{i=1}^n X_i \leq y) \approx \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2/n}\right) dy.$$

The approximation error can be bounded by using the *Berry-Essene* Theorems. See Serfling [62] for details.

3.1.3 CHI-SQUARE

The (central) **Chi-square** density with k degrees of freedom (df) is of the form:

$$f_\theta(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}, \quad x > 0, \tag{17}$$

where $\theta = k$, a positive integer. Here $\Gamma(u)$ denotes the Gamma function,

$$\Gamma(u) = \int_0^\infty x^{u-1} e^{-x} dx,$$

For n integer valued $\Gamma(n+1) = n! = n(n-1)\ldots 1$ and $\Gamma(n+1/2) = \frac{(2n-1)(2n-3)\ldots 5\cdot 3\cdot 1}{2^n}\sqrt{\pi}$.

If $Z_i \sim \mathcal{N}(0,1)$ are i.i.d., $i = 1, \ldots, n$, then $X = \sum_{i=1}^n Z_i^2$ is distributed as Chi-square with n degrees of freedom (df). Our shorthand notation for this is

$$\sum_{i=1}^n [\mathcal{N}(0,1)]^2 = \chi_n. \tag{18}$$

This characterization of a Chi square r.v. is sometimes called a *stochastic representation* since it is defined via operations on other r.v.s. The fact that (17) is the density of a sum of squares of independent $\mathcal{N}(0,1)$'s is easily derived. Start with the density function $f(z) = e^{-z^2/2}/\sqrt{2\pi}$ of a standard Gaussian random variable Z. Using the relation $(\sqrt{2\pi}\sigma)^{-1}\int_{-\infty}^{\infty} e^{-u^2/(2\sigma^2)} du = 1$, the characteristic function of Z^2 is simply found as $\Phi_{Z^2}(u) = E[e^{juZ^2}] = (1+j2u)^{-1/2}$. Applying the summation-convolution theorem for independent r.v.s Y_i, $\Phi_{\sum Y_i}(u) = \prod \Phi_{Y_i}(u)$, we obtain $\Phi_{\sum_{i=1}^n Z_i^2}(u) = (1+j2u)^{-n/2}$. Finally, using a table of Fourier transform relations, identify (17) as the inverse fourier transform of $\Phi_{\sum_{i=1}^n Z_i^2}(u)$.

Some useful properties of the Chi-square random variable are as follows:

* $E[\chi_n] = n, \quad \text{var}(\chi_n) = 2n$

* Asymptotic relation for large n:

$$\chi_n = \sqrt{2n}\mathcal{N}(0,1) + n$$

* χ_2 an exponential r.v. with mean 2, i.e. $X = \chi_2$ is a non-negative r.v. with probability density $f(x) = \frac{1}{2}e^{-x/2}$.

* $\sqrt{\chi_2}$ is a Rayleigh distributed random variable.

3.1.4 GAMMA

The Gamma density function is

$$f_{\underline{\theta}}(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, \quad x > 0,$$

where $\underline{\theta}$ denotes the pair of parameters (λ, r), $\lambda, r > 0$. Let $\{Y_i\}_{i=1}^n$ be i.i.d. exponentially distributed random variables with mean $1/\lambda$, specifically Y_i has density

$$f_\lambda(y) = \lambda e^{-\lambda y}, \quad y > 0.$$

Then the sum $X = \sum_{i=1}^n Y_i$ has a Gamma density $f_{(\lambda,n)}$. Other useful properties of a Gamma distributed random variable X with parameters $\underline{\theta} = (\lambda, r)$ include:

* $E_{\underline{\theta}}[X] = r/\lambda$

* $\text{var}_{\underline{\theta}}(X) = r/\lambda^2$

* The Chi-square distribution with k df is a special case of the Gamma distribution obtained by setting Gamma parameters as follows: $\lambda = 1/2$ and $r = k/2$.

3.1.5 NON-CENTRAL CHI SQUARE

The sum of squares of independent Gaussian r.v.s with unit variances but non-zero means is called a **non-central Chi-square** r.v. Specifically, if $Z_i \sim \mathcal{N}(\mu_i, 1)$ are independent, $i = 1, \ldots, n$, then $X = \sum_{i=1}^n Z_i^2$ is distributed as non-central Chi-square with n df and non-centrality parameter $\delta = \sum_{i=1}^n \mu_i^2$. In our shorthand we write

$$\sum_{i=1}^n [\mathcal{N}(0,1) + \mu_i]^2 = \sum_{i=1}^n [\mathcal{N}(\mu_i, 1)]^2 = \chi_{n,\delta}. \tag{19}$$

The non-central Chi-square density has no simple expression of closed form. There are some useful asymptotic relations, however:

* $E[\chi_{n,\delta}] = n + \delta, \quad \text{var}(\chi_{n,\delta}) = 2(n + 2\delta)$

* $\sqrt{\chi_{2,\mu_1^2+\mu_2^2}}$ is a Rician r.v.

3.1.6 CHI-SQUARE MIXTURE

The distribution of the sum of squares of independent Gaussian r.v.s with zero mean but different variances is not closed form either. However, many statisticians have studied and tabulated the distribution of a weighted sum of squares of i.i.d. standard Gaussian r.v.s $Z_1, \ldots, Z_n$, $Z_i \sim \mathcal{N}(0,1)$. Specifically, the following has a (central) **Chi-square mixture** (also known as the Chi-bar square [30]) with n degrees of freedom and mixture parameter $\underline{c} = [c_1, \ldots, c_n]^T$, $c_i \geq 0$:

$$\sum_{i=1}^{n} \frac{c_i}{\sum_j c_j} Z_i^2 = \overline{\chi}_{n,\underline{c}}$$

An asymptotic relation of interest to us will be:

* $E[\overline{\chi}_{n,\underline{c}}] = 1$, $\quad$, $\mathrm{var}(\overline{\chi}_{n,\underline{c}}) = 2\sum_{i=1}^{N}\left(\frac{c_i}{\sum_j c_i}\right)^2$

Furthermore, there is an obvious a special case where the Chi square mixture reduces to a scaled (central) Chi square: $\overline{\chi}_{n,c\underline{1}} = \frac{1}{n}\chi_n$ for any $c \neq 0$.

3.1.7 STUDENT-T

For $Z \sim \mathcal{N}(0,1)$ and $Y \sim \chi_n$ independent r.v.s the ratio $X = Z/\sqrt{Y/n}$ is called a **Student-t** r.v. with n degrees of freedom, denoted $\mathcal{T}_n$. Or in our shorthand notation:

$$\frac{\mathcal{N}(0,1)}{\sqrt{\chi_n/n}} = \mathcal{T}_n.$$

The density of $\mathcal{T}_n$ is the Student-t density with n df and has the form

$$f_{\underline{\theta}}(x) = \frac{\Gamma([n+1]/2)}{\Gamma(n/2)} \frac{1}{\sqrt{n\pi}} \frac{1}{(1+x^2/n)^{(n+1)/2}}, \quad x \in \mathbb{R},$$

where $\underline{\theta} = n$ is a positive integer. Properties of interest to us are:

* $E[\mathcal{T}_n] = 0$ $(n > 1)$, $\mathrm{var}(\mathcal{T}_n) = \frac{n}{n-2}$ $(n > 2)$

* Asymptotic relation for large n:

$$\mathcal{T}_n \approx \mathcal{N}(0,1).$$

For $n = 1$ the mean of $\mathcal{T}_n$ does not exist and for $n \leq 2$ its variance is infinite.

3.1.8 FISHER-F

For $U \sim \chi_m$ and $V \sim \chi_n$ independent r.v.s the ratio $X = (U/m)/(V/n)$ is called a **Fisher-F** r.v. with m, n degrees of freedom, or in shorthand:

$$\frac{\chi_m/m}{\chi_n/n} = \mathcal{F}_{m,n}.$$

The Fisher-F density with m and n df is defined as

$$f_{\underline{\theta}}(x) = \frac{\Gamma([m+n]/2)}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{n}\right)^{m/2} \frac{x^{(m-2)/2}}{(1+\frac{m}{n}x)^{(m+n)/2}}, \quad x > 0$$

where $\underline{\theta} = [m, n]$ is a pair of positive integers. It should be noted that moments $E[X^k]$ of order greater than $k = n/2$ do not exist. A useful asymptotic relation for n large and $n \gg m$ is

$$\mathcal{F}_{m,n} \approx \chi_m.$$

3.1.9 CAUCHY

The ratio of independent $\mathcal{N}(0, 1)$ r.v.'s U and V is called a standard Cauchy r.v.

$$X = U/V \sim \mathcal{C}(0, 1).$$

It's density has the form

$$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2} \quad x \in \mathbb{R}$$

. If $\underline{\theta} = [\mu, \sigma]$ are location and scale parameters $(\sigma > 0)$ $f_\theta(x) = f((x - \mu)/\sigma)$ is a translated and scaled version of the standard Cauchy density denoted $\mathcal{C}(\mu, \sigma^2)$. Some properties of note: (1) the Cauchy distribution has no moments of any (positive) integer order; and (2) the Cauchy distribution is the same as a Student-t distribution with 1 d.f.

3.1.10 BETA

For $U \sim \chi_m$ and $V \sim \chi_n$ independent Chi-square r.v.s with m and n df, respectively, the ratio $X = U/(U + V)$ has a **Beta** distribution, or in shorthand

$$\frac{\chi_m}{\chi_m + \chi_n} = \mathcal{B}(m/2, n/2)$$

where $\mathcal{B}(p, q)$ is a r.v. with Beta density having paramaters $\underline{\theta} = [p, q]$. The Beta density has the form

$$f_{\underline{\theta}}(x) = \frac{1}{\beta_{r,t}} x^{r-1}(1 - x)^{t-1}, \quad x \in [0, 1]$$

where $\underline{\theta} = [r, t]$ and $r, t > 0$. Here $\beta_{r,t}$ is the Beta function:

$$\beta_{r,t} = \int_0^1 x^{r-1}(1 - x)^{t-1} dx = \frac{\Gamma(r)\Gamma(t)}{\Gamma(r + t)}.$$

Some useful properties:

* The special case of $m = n = 1$ gives rise to X an **arcsin** distributed r.v.

* $E_{\underline{\theta}}[\mathcal{B}(p, q)] = p/(p + q)$

* $\text{var}_{\underline{\theta}}(\mathcal{B}(p, q)) = pq/((p + q + 1)(p + q)^2)$

3.2 REPRODUCING DISTRIBUTIONS

A random variable X is said to have a *reproducing distribution* if the sum of two independent realizations, say X_1 and X_2, of X have the same distribution, possibly with different parameter values, as X. A Gaussian r.v. has a reproducing distribution:

$$\mathcal{N}(\mu_1, \sigma_1^2) + \mathcal{N}(\mu_2, \sigma_2^2) = \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2),$$

which follows from the fact that the convolution of two Gaussian density functions is a Gaussian density function [48]. Noting the stochastic representations (18) and (19) of the Chi square and non-central Chi square distributions, respectively, it is obvious that they are reproducing distributions:

*$\chi_n + \chi_m = \chi_{m+n}$, if χ_m, χ_n are independent.

*$\chi_{m,\delta_1} + \chi_{n,\delta_2} = \chi_{m+n,\delta_1+\delta_2}$, if χ_{m,δ_1}, χ_{n,δ_2} are independent.

The Chi square mixture, Fisher-F, and Student-t are not reproducing densities.

3.3 FISHER-COCHRAN THEOREM

This result gives a very useful tool for finding the distribution of quadratic forms of Gaussian random variables. A more general result that covers the joint distribution of quadratic forms is given in [57].

Theorem 1 *Let $\underline{X} = [X_1, \ldots, X_n]^T$ be a vector of iid. $\mathcal{N}(0,1)$ rv's and let $\mathbf{A}$ be a symmetric idempotent matrix ($\mathbf{AA} = \mathbf{A}$) of rank p. Then*

$$\underline{X}^T \mathbf{A} \underline{X} = \chi_p$$

A simple proof is given below.

Proof: Let $\mathbf{A} = \mathbf{U\Lambda U}^T$ be the eigendecomposition of $\mathbf{A}$. Then

* All eigenvalues λ_i of $\mathbf{A}$ are either 0 or 1

$$\begin{aligned}\mathbf{AA} &= \mathbf{U\Lambda}\underbrace{\mathbf{U}^T\mathbf{U}}_{=\mathbf{I}}\mathbf{\Lambda U}^T \\ &= \mathbf{U\Lambda}^2\mathbf{U}^T = \mathbf{U\Lambda U}^T\end{aligned}$$

and therefore

$$\begin{aligned}\underline{X}^T\mathbf{A}\underline{X} &= \underline{X}^T\mathbf{U\Lambda}\underbrace{\mathbf{U}^T\underline{X}}_{\underline{Z}=\mathcal{N}_n(0,\mathbf{I})} \\ &= \sum_{i=1}^n \lambda_i Z_i^2 = \sum_{i=1}^p [\mathcal{N}(0,1)]^2\end{aligned}$$

$\diamond$

3.4 SAMPLE MEAN AND SAMPLE VARIANCE

Let X_i's be i.i.d. $\mathcal{N}(\mu, \sigma^2)$ r.v.'s. The sample mean and sample variance respectively approximate the location μ and spread σ of the population.

* Sample mean: $\overline{X} = n^{-1}\sum_{i=1}^n X_i$

* Sample variance: $\mathsf{s}^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$

In the Gaussian case the joint distribution of the sample mean and variance can be specified.

(1). $\overline{X} \ = \ \mathcal{N}(\mu, \sigma^2/n)$

(2). $\mathsf{s}^2 \ = \ \frac{\sigma^2}{n-1} \chi_{n-1}$

(3). $\overline{X}$ and s^2 are independent rv's.

These results imply that a weighted ratio of sample mean and sample variance is distributed as Student t.

$$\frac{\overline{X} - \mu}{\mathsf{s}/\sqrt{n}} = \mathcal{T}_{n-1}.$$

Proof of assertions (2) and (3): In view of the representation (13), it suffices consider the the case of a standard Gaussian sample: $\mu = 0$ and $\sigma = 1$.

First we show that the sample mean and the sample variance are independent random variables. Define the vector of random variables $\underline{Y} = [Y_1, \ldots, Y_n]^T$ as follows. First define

$$Y_1 = \sqrt{n}\overline{X} = \underline{h}_1^T \underline{X},$$

where

$$\underline{h}_1 = [1/\sqrt{n}, \ldots, 1/\sqrt{n}]^T.$$

Note that $\underline{h}_1$ has unit norm. Next apply the Gramm-Schmidt orthonormalization procedure of Sec. 2.3.6 to complete the basis with respect to $\underline{h}_1$. This generates $n-1$ vectors $\underline{h}_2, \ldots, \underline{h}_n$ that are orthonormal, mutually orthogonal, and orthogonal to $\underline{h}_1$. The random vector $\underline{Y}$ is now defined as

$$\underline{Y} = \mathbf{H}^T \underline{X}$$

where $\mathbf{H} = [\underline{h}_1, \ldots, \underline{h}_n]$ is an $n \times n$ orthogonal matrix.

Since, $\mathbf{X} = \mathbf{HY}$, the orthogonality of $\mathbf{H}$ implies the following properties

1. The Y_i's are zero mean unit variance independent Gaussian random variables: $\underline{Y} \ \sim \ \mathcal{N}_n(\underline{0}, \mathbf{I})$
2. $\underline{Y}^T \underline{Y} = \underline{X}^T \underline{X}$

As $\underline{Y}_1 \ = \ \sqrt{n}\overline{X}$ Property 1 implies that $\overline{X}$ is independent of $Y_2, \ldots, Y_n$. Furthermore, using the equivalence:

$$\sum_{i=1}^n (X_i - \overline{X})^2 = \sum_{i=1}^n X_i^2 - n(\overline{X})^2,$$

Property 2 and the definition of Y_1 imply that

$$\sum_{i=1}^n (X_i - \overline{X})^2 = \sum_{i=1}^n Y_i^2 - Y_1^2 = Y_2^2 + \cdots + Y_n^2, \tag{20}$$

that is, the sample variance is only a function of $Y_2, \ldots, Y_n$ and is therefore independent of $Y_1 =$ the sample mean.

Furthermore, as $Y_2, \ldots, Y_n$ are independent $\mathcal{N}(0,1)$ random variables, the representation (20) implies that the (normalized) sample variance has a Chi-square distribution with $n-1$ degrees of freedom.

This completes the proof of assertions (2) and (3). $\diamond$

The Chi-square property in assertion (3) can also be shown directly using the Fisher-Cochran theorem (Thm. 1). Note that the normalized sample variance on the extreme left of the equalities (20) can be expressed as a quadratic form

$$\begin{aligned}[\underline{X} - \underline{1}\overline{X}]^T[\underline{X} - \underline{1}\overline{X}] &= \underline{X}^T \underbrace{[\mathbf{I} - \underline{1}\underline{1}^T\frac{1}{n}]}_{\text{idempotent}} [\mathbf{I} - \underline{1}\underline{1}^T\frac{1}{n}]\underline{X}^T \\ &= \underline{X}^T \underbrace{[\mathbf{I} - \underline{1}\underline{1}^T\frac{1}{n}]}_{\text{orth. proj.}} \underline{X}\end{aligned}$$

where $\underline{1} = [1, \ldots, 1]^T$. Observe: since $\text{rank}[\mathbf{I} - \underline{1}\underline{1}^T\frac{1}{n}] = n - 1$, we have that $[\underline{X} - \underline{1}\overline{X}]^T[\underline{X} - \underline{1}\overline{X}] = (n-1)\,\mathsf{s}^2$ is χ_{n-1}.

3.5 SUFFICIENT STATISTICS

Many detection/estimation/classification problems have the following common structure. A continuous time waveform $\{x(t) : t \in \mathbb{R}\}$ is measured at n time instants $t_1, \ldots, t_n$ producing the vector

$$\underline{x} = [x_1, \ldots, x_n]^T,$$

where $x_i = x(t_i)$. The vector $\underline{x}$ is modelled as a realization of a random vector $\underline{X}$ with a joint distribution which is of known form but depends on a handful (p) of unknown parameters $\underline{\theta} = [\theta_1, \ldots, \theta_p]^T$.

More concisely:

* $\underline{X} = [X_1, \ldots, X_n]^T$, $X_i = X(t_i)$, is a vector of random measurements or observations taken over the course of the experiment

* $\mathcal{X}$ is sample or measurement space of realizations $\underline{x}$ of $\underline{X}$

* $\mathcal{B}$ is the event space induced by $\underline{X}$, e.g., the Borel subsets of $\mathbb{R}^n$

* $\underline{\theta} \in \Theta$ is an unknown parameter vector of interest

* Θ is parameter space for the experiment

* $P_{\underline{\theta}}$ is a probability measure on $\mathcal{B}$ for given $\underline{\theta}$. $\{P_{\underline{\theta}}\}_{\underline{\theta}\in\Theta}$ is called the *statistical model* for the experiment.

The probability model induces the joint cumulative distribution function j.c.d.f. associated with $\underline{X}$

$$F_{\underline{X}}(\underline{x}; \theta) = P_{\underline{\theta}}(X_1 \le x_1, \ldots, X_n \le x_n),$$

which is assumed to be known for any $\underline{\theta} \in \Theta$. When $\underline{X}$ is a continuous random variable the j.c.d.f. is specified by the joint probability density function (j.p.d.f.) that we will write in several different ways, depending on the context: $f_{\underline{\theta}}(\underline{x})$ or $f(\underline{x}; \underline{\theta})$, or, when we need to explicitly call out the r.v. $\underline{X}$, $f_X(\underline{x}; \underline{\theta})$. We will denote by $E_{\underline{\theta}}[Z]$ the statistical expectation of a random variable Z with respect to the j.p.d.f. $f_Z(z; \underline{\theta})$

$$E_{\underline{\theta}}[Z] = \int z f_Z(z; \underline{\theta}) dz.$$

The family of functions $\{f(\underline{x}; \underline{\theta})\}_{\underline{x}\in\mathcal{X}, \underline{\theta}\in\Theta}$ then defines the statistical model for the experiment.

The general objective of statistical inference can now be stated. Given a realization $\underline{x}$ of $\underline{X}$ infer properties of $\underline{\theta}$ knowing only the parametric form of the statistical model. Thus we will want to come up with a function, called an inference function, which maps $\underline{X}$ to subsets of the parameter space, e.g., an estimator, classifier, or detector for $\underline{\theta}$. As we will see later there are many ways to design inference functions but a more fundamental question is: are there any general properties that good inference functions should have? One such property is that the inference function only need depend on the n-dimensional data vector $\underline{X}$ through a lower dimensional version of the data called a *sufficient statistic*.

3.5.1 SUFFICIENT STATISTICS AND THE REDUCTION RATIO

First we define a statistic as any function $T = T(\underline{X})$ of the data (actually, for T to be a valid random variable derived from $\underline{X}$ it must be a *measurable* function, but this theoretical technicality is beyond our scope here).

There is a nice interpretation of a statistic in terms of its memory storage requirements. Assume that you have a special computer that can store any one of the time samples in $\underline{X} = [X_1, \ldots, X_n]$, $X_k = X(t_k)$ say, in a "byte" of storage space and the time stamp t_k in another "byte" of storage space. Any non-invertible function T, e.g., which maps $\mathbb{R}^n$ to a lower dimensional space $\mathbb{R}^m$, can be viewed as a dimensionality reduction on the data sample. We can quantify the amount of reduction achieved by T by defining the reduction ratio (RR):

$$RR = \frac{\text{\# bytes of storage required for } T(\underline{X})}{\text{\# bytes of storage required for } \underline{X}}$$

This ratio is a measure of the amount of data compression induced by a specific transformation T. The number of bytes required to store $\underline{X}$ with its time stamps is:

$$\text{\# bytes}\{\underline{X}\} = \text{\# bytes}[X_1, \ldots, X_n]^T = \text{\# bytes}\{\text{timestamps}\} + \text{\# bytes}\{values\} = 2n$$

Consider the following examples:

Define $X_{(i)} =$ as the i-th largest element of $\underline{X}$. The $X_{(i)}$'s satisfy: $X_{(1)} \geq X_{(2)} \geq \ldots \geq X_{(n)}$ and are nothing more than a convenient reordering of the data sample $X_1, \ldots, X_n$. The $X_{(i)}$'s are called the *rank ordered statistics* and do not carry time stamp information. The following table illustrates the reduction ratio for some interesting cases

Statistic used	**Meaning in plain english**	**Reduction ratio**
$T(\underline{X}) = [X_1, \ldots, X_n]^T$,	entire data sample	RR $= 1$
$T(\underline{X}) = [X_{(1)}, \ldots, X_{(n)}]^T$,	rank ordered sample	RR $= 1/2$
$T(\underline{X}) = \overline{X}$,	sample mean	RR $= 1/(2n)$
$T(\underline{X}) = [\overline{X}, \mathsf{s}^2]^T$,	sample mean and variance	RR $= 1/n$

A natural question is: what is the maximal reduction ratio one can get away with without loss of information about $\underline{\theta}$? The answer is: the ratio obtained by compression to a quantity called a *minimal sufficient statistic*. But we are getting ahead of ourselves. We first need to define a plain old sufficient statistic.

3.5.2 DEFINITION OF SUFFICIENCY

Here is a warm up before making a precise definition of sufficiency. $T = T(\underline{X})$ is a **sufficient statistic** (SS) for a parameter $\underline{\theta}$ if it captures all the information in the data sample useful for inferring the value of $\underline{\theta}$. To put it another way: once you have computed a sufficient statistic you can store it and throw away the original sample since keeping it around would not add any useful information.

More concretely, let $\underline{X}$ have a cumulative distribution function (CDF) $F_{\underline{X}}(\underline{x};\underline{\theta})$ depending on $\underline{\theta}$. A statistic $T = T(\underline{X})$ is said to be sufficient for $\underline{\theta}$ if the conditional CDF of $\underline{X}$ given $T = t$ is not a function of $\underline{\theta}$, i.e.,

$$F_{\underline{X}|T}(\underline{x}|T = t, \underline{\theta}) = G(\underline{x}, t), \tag{21}$$

where G is a function that does not depend on $\underline{\theta}$.

Specializing to a discrete valued $\underline{X}$ with probability mass function $p_{\underline{\theta}}(\underline{x}) = P_{\underline{\theta}}(\underline{X} = \underline{x})$, a statistic $T = T(\underline{X})$ is sufficient for $\underline{\theta}$ if

$$P_{\underline{\theta}}(\underline{X} = \underline{x}|T = t) = G(\underline{x}, t). \tag{22}$$

For a continuous r.v. $\underline{X}$ with pdf $f(\underline{x};\underline{\theta})$, the condition (21) for T to be a sufficient statistic (SS) becomes:

$$f_{\underline{X}|T}(\underline{x}|t;\underline{\theta}) = G(\underline{x}, t). \tag{23}$$

Sometimes the only sufficient statistics are vector statistics, e.g. $T(\underline{X}) = \underline{T}(\underline{X}) = [T_1(\underline{X}), \ldots, T_K(\underline{X})]^T$. In this case we say that the T_k's are *jointly sufficient* for $\underline{\theta}$

The definition (21) is often difficult to use since it involves derivation of the conditional distribution of $\underline{X}$ given T. When the random variable $\underline{X}$ is discrete or continuous a simpler way to verify sufficiency is through the Fisher factorization (FF) property [57]

Fisher factorization (FF): $T = T(\underline{X})$ is a sufficient statistic for $\underline{\theta}$ if the probability density $f_{\underline{X}}(\underline{x};\underline{\theta})$ of $\underline{X}$ has the representation

$$f_{\underline{X}}(\underline{x};\underline{\theta}) = g(T, \underline{\theta})\ h(\underline{x}), \tag{24}$$

for some non-negative functions g and h. The FF can be taken as the operational definition of a sufficient statistic T. An important implication of the Fisher Factorization is that when the density function of a sample $\underline{X}$ satisfies (24) then the density $f_T(t;\underline{\theta})$ of the sufficient statistic T is equal to $g(t, \underline{\theta})$ up to a $\underline{\theta}$-independent constant $q(t)$ (see exercises at end of this chapter):

$$f_T(t;\underline{\theta}) = g(t, \underline{\theta}) q(t).$$

Examples of sufficient statistics:

Example 1 *Entire sample*

$\underline{X} = [X_1, \ldots, X_n]^T$ is sufficient but not very interesting

Example 2 *Rank ordered sample*

$X_{(1)}, \ldots, X_{(n)}$ is sufficient when X_i's i.i.d.

Proof: Since X_i's are i.i.d., the joint pdf is

$$f_\theta(x_1, \ldots, x_n) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n f_\theta(x_{(i)}).$$

Hence sufficiency of the rank ordered sample $X_{(1)}, \ldots, X_{(n)}$ follows from Fisher factorization.

Example 3 *Binary likelihood ratios*

Let $\underline{\theta}$ take on only two possible values $\underline{\theta}_0$ and $\underline{\theta}_1$, e.g., a bit taking on the values "0" or "1" in a communication link. Then, as $f(\underline{x}; \underline{\theta})$ can only be $f(\underline{x}; \underline{\theta}_0)$ or $f(\underline{x}; \underline{\theta}_1)$, we can reindex the pdf as $f(\underline{x}; \theta)$ with the scalar parameter $\theta \in \Theta = \{0, 1\}$. This gives the binary decision problem: "decide between $\theta = 0$ versus $\theta = 1$." If it exists, i.e. it is finite for all values of $\underline{X}$, the "likelihood ratio" $\Lambda(\underline{X}) = f_1(\underline{X})/f_0(\underline{X})$ is sufficient for θ, where $f_1(\underline{x}) \stackrel{\text{def}}{=} f(\underline{x}; 1)$ and $f_0(\underline{x}) \stackrel{\text{def}}{=} f(\underline{x}; 0)$.

Proof: Express $f_\theta(\underline{X})$ as function of θ, f_0, f_1, factor out f_0, identify Λ, and invoke FF

$$\begin{aligned} f_{\underline{\theta}}(\underline{X}) &= \theta f_1(\underline{X}) + (1-\theta) f_0(\underline{X}) \\ &= \left(\underbrace{\theta \Lambda(\underline{X}) + (1-\theta)}_{g(T,\theta)} \right) \underbrace{f_0(\underline{X})}_{h(\underline{X})}. \end{aligned}$$

$\diamond$

Therefore to discriminate between two values $\underline{\theta}_1$ and $\underline{\theta}_2$ of a parameter vector $\underline{\theta}$ we can throw away all data except for the scalar sufficient statistic $T = \Lambda(\underline{X})$

Example 4 *Discrete likelihood ratios*

Let $\Theta = \{\underline{\theta}_1, \ldots, \underline{\theta}_p\}$ and assume that the vector of $p-1$ *likelihood ratios*

$$\underline{T}(\underline{X}) = \left[\frac{f_{\theta_1}(\underline{X})}{f_{\theta_p}(\underline{X})}, \ldots, \frac{f_{\theta_{p-1}}(\underline{X})}{f_{\theta_p}(\underline{X})} \right]^T = [\Lambda_1(\underline{X}), \ldots, \Lambda_{p-1}(\underline{X})]^T$$

is finite for all $\underline{X}$. Then this vector is sufficient for θ. An equivalent way to express this vector is as the sequence $\{\Lambda_\theta(\underline{X})\}_{\theta \in \Theta} = \Lambda_1(\underline{X}), \ldots, \Lambda_{p-1}(\underline{X})$, and this is called the *likelihood trajectory over* θ.

Proof

Define the $p-1$ element selector vector $\underline{u}_\theta = \underline{e}_k$ when $\theta = \theta_k$, $k = 1, \ldots, p-1$ (recall that $\underline{e}_k = [0, \ldots, 0, 1, 0, \ldots 0]^T$ is the k-th column of the $(p-1) \times (p-1)$ identity matrix). Now for any $\theta \in \Theta$ we can represent the j.p.d.f. as

$$f_\theta(\underline{x}) = \underbrace{\underline{u}_\theta^T \underline{T}}_{g(\underline{T},\theta)} \; \underbrace{f_{\theta_p}(\underline{x})}_{h(\underline{x})},$$

which establishes sufficiency by the FF. $\diamond$

Example 5 *Likelihood ratio trajectory*

When Θ is a set of scalar parameters θ the likelihood ratio trajectory over Θ

$$\Lambda(\underline{X}) = \left\{ \frac{f_\theta(\underline{X})}{f_{\theta_0}(\underline{X})} \right\}_{\theta\in\Theta}, \tag{25}$$

is sufficient for θ. Here θ_0 is an arbitrary reference point in Θ for which the trajectory is finite for all $\underline{X}$. When θ is not a scalar (25) becomes a likelihood ratio surface, which is also a sufficient statistic.

3.5.3 MINIMAL SUFFICIENCY

What is the maximum possible amount of reduction one can apply to the data sample without losing information concerning how the model depends on $\underline{\theta}$? The answer to this question lies in the notion of a minimal sufficient statistic. Such statistics cannot be reduced any further without loss in information. In other words, any other sufficient statistic can be reduced down to a minimal sufficient statistic without information loss. Since reduction of a statistic is accomplished by applying a functional transformation we have the formal definition.

Definition: T_{min} is a minimal sufficient statistic if it can be obtained from any other sufficient statistic T by applying a functional transformation to T. Equivalently, if T is any sufficient statistic there exists a function q such that $T_{min} = q(T)$.

Minimal sufficient statistics are not unique: if $T_{\min}$ is minimal sufficient $h(T_{min})$ is also minimal sufficient if h is any invertible function. Minimal sufficient statistics can be found in a variety of ways [48, 7, 41]. One way is to find a *complete sufficient statistic*; under broad conditions this statistic will also be minimal [41]. A sufficient statistic T is complete if

$$E_{\underline{\theta}}[g(T)] = 0, \quad \text{for all } \underline{\theta} \in \Theta$$

implies that the function g is identically zero, i.e., $g(t) = 0$ for all values of t.

To see that a completeness implies minimality we can adapt the proof of Scharf in [60]. Let M be a minimal sufficient statistic and let C be complete sufficient statistic. As M is minimal it is a function of C. Therefore $g(C) \stackrel{\text{def}}{=} C - E_\theta[C|M]$ is a function of C since the conditional expectation $E_\theta[X|M]$ is a function of M. Since, obviously, $E_\theta[g(C)] = 0$ for all θ and C is complete, $C = E_\theta[C|M]$ for all θ. Thus C is minimal since it is a function of M which is a function of any other sufficient statistic. In other words, C inherits minimality from M.

Another way to find a minimal sufficient statistic is through reduction of the data to the likelihood ratio surface.

As in Example 5, assume that there exists a reference point $\underline{\theta}_o \in \Theta$ such that the following likelihood-ratio function is finite for all $\underline{x} \in \mathcal{X}$ and all $\theta \in \Theta$

$$\Lambda_{\underline{\theta}}(\underline{x}) = \frac{f_{\underline{\theta}}(\underline{x})}{f_{\underline{\theta}_o}(\underline{x})}.$$

For given $\underline{x}$ let $\Lambda(\underline{x})$ denote the set of likelihood ratios (a likelihood ratio trajectory or surface)

$$\Lambda(\underline{x}) = \{\Lambda_{\underline{\theta}}(\underline{x})\}_{\underline{\theta}\in\Theta}.$$

Definition 1 *We say that a (θ-independent) function of $\underline{x}$, denoted $\tau = \tau(\underline{x})$, indexes the likelihood ratios Λ when both*

1. $\Lambda(\underline{x}) = \Lambda(\tau)$, *i.e., Λ only depends on $\underline{x}$ through $\tau = \tau(\underline{x})$.*
2. $\Lambda(\tau) = \Lambda(\tau')$ *implies $\tau = \tau'$, i.e., the mapping $\tau \to \Lambda(\tau)$ is invertible.*

Condition 1 is an equivalent way of stating that $\tau(\underline{X})$ is a sufficient statistic for $\underline{\theta}$.

Theorem:If $\tau = \tau(\underline{x})$ indexes the likelihood ratios $\Lambda(\underline{x})$ then $T_{min} = \tau(\underline{X})$ is minimally sufficient for $\underline{\theta}$.

Proof:

We prove this only for the case that $\underline{X}$ is a continuous r.v. First, condition 1 in Definition 1 implies that $\tau = \tau(\underline{X})$ is a sufficient statistic. To see this use FF and the definition of the likelihood ratios to see that $\Lambda(\underline{x}) = \Lambda(\tau)$ implies: $f_{\underline{\theta}}(\underline{X}) = \Lambda_{\underline{\theta}}(\tau) f_{\underline{\theta}_o}(\underline{X}) = g(\tau; \underline{\theta}) h(\underline{x})$. Second, let T be any sufficient statistic. Then, again by FF, $f_{\underline{\theta}}(\underline{x}) = g(T, \underline{\theta})\ h(\underline{x})$ and thus

$$\Lambda(\tau) = \left\{ \frac{f_{\underline{\theta}}(\underline{X})}{f_{\underline{\theta}_o}(\underline{X})} \right\}_{\underline{\theta} \in \Theta} = \left\{ \frac{g(T, \underline{\theta})}{g(T, \underline{\theta}_o)} \right\}_{\underline{\theta} \in \Theta}.$$

so we conclude that $\Lambda(\tau)$ is a function of T. But by condition 2 in Definition 1 the mapping $\tau \to \Lambda(\tau)$ is invertible and thus τ is itself a function of T. $\diamond$

Another important concept in practical applications is that of finite dimensionality of a sufficient statistic.

Definition: a sufficient statistic $T(\underline{X})$ is said to be **finite dimensional** if its dimension is not a function of the number of data samples n.

Frequently, but not always (see Cauchy example below), minimal sufficient statistics are finite dimensional.

Example 6 *Minimal sufficient statistic for mean of Gaussian density.*

Assume $X \sim \mathcal{N}(\mu, \sigma^2)$ where σ^2 is known. Find a minimal sufficient statistic for $\theta = \mu$ given the iid sample $\underline{X} = [X_1, \ldots, X_n]^T$.

Solution: the j.p.d.f. is

$$\begin{aligned}
f_\theta(\underline{x}) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2} \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right)} \\
&= \underbrace{e^{-\frac{n\mu^2}{2\sigma^2}}\, e^{\mu/\sigma^2 \overbrace{\sum_{i=1}^n x_i}^{T(\underline{x})}}}_{g(\underline{T},\theta)}\ \underbrace{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-1/(2\sigma^2)\sum_{i=1}^n x_i^2}}_{h(\underline{x})}
\end{aligned}$$

Thus by FF

$$T = \sum_{i=1}^n X_i$$

is a sufficient statistic for μ. Furthermore, as $q(T) = n^{-1}T$ is a 1-1 function of T

$$\underline{S} = \overline{X}$$

is an equivalent sufficient statistic.

Next we show that the sample mean is in fact minimal sufficient by showing that it indexes the likelihood ratio trajectory $\Lambda(\underline{x}) = \{\Lambda_\theta(\underline{x})\}_{\theta\in\Theta}$, with $\theta = \mu$, $\Theta = \mathbb{R}$. Select the reference point $\theta_o = \mu_o = 0$ to obtain:

$$\Lambda_\mu(\underline{x}) = \frac{f_\mu(\underline{x})}{f_0(\underline{x})} = \exp\left(\mu/\sigma^2 \sum_{i=1}^n x_i - \tfrac{1}{2} n\mu^2/\sigma^2\right).$$

Identifying $\tau = \sum_{i=1}^n x_i$, condition 1 in Definition 1 is obviously satisfied since $\Lambda_\mu(\underline{x}) = \Lambda_\mu(\sum x_i)$ (we already knew this since we showed that $\sum_{i=1}^n X_i$ was a sufficient statistic). Condition 2 in Definition 1 follows since $\Lambda_\mu(\sum x_i)$ is an invertible function of $\sum x_i$ for any non-zero value of μ (summation limits omitted for clarity). Therefore the sample mean indexes the trajectories, and is minimal sufficient.

Example 7 *Minimal sufficient statistics for mean and variance of Gaussian density.*

Assume $X \sim \mathcal{N}(\mu, \sigma^2)$ where both μ and σ^2 are unknown. Find a minimal sufficient statistic for $\underline{\theta} = [\mu, \sigma^2]^T$ given the iid sample $\underline{X} = [X_1, \ldots, X_n]^T$.

Solution:

$$\begin{aligned}
f_{\underline{\theta}}(\underline{x}) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2} \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right)} \\
&= \underbrace{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{n\mu^2}{2\sigma^2}}\, e^{[\mu/\sigma^2,\ -1/(2\sigma^2)]\overbrace{\left[\sum_{i=1}^n x_i,\ \sum_{i=1}^n x_i^2\right]^T}^{\underline{T}(\underline{x})}}}_{g(\underline{T},\underline{\theta})} \underbrace{1}_{h(\underline{x})}
\end{aligned}$$

Thus

$$\underline{T} = \left[\underbrace{\sum_{i=1}^n X_i}_{T_1},\ \underbrace{\sum_{i=1}^n X_i^2}_{T_2}\right]$$

is a (jointly) sufficient statistic for μ, σ^2. Furthermore, as $q(\underline{T}) = [n^{-1}T_1, (n-1)^{-1}(T_2 - T_1^2)]$ is a 1-1 function of $\underline{T}$ ($\underline{T} = [T_1, T_2]^T$)

$$\underline{S} = \left[\overline{X},\ \mathsf{s}^2\right]$$

is an equivalent sufficient statistic.

Similarly to Example 6, we can show minimal sufficiency of this statistic by showing that it indexes the likelihood ratio surface $\{\Lambda_{\underline{\theta}}(\underline{X})\}_{\underline{\theta}\in\Theta}$, with $\theta = [\mu, \sigma^2]$, $\Theta = \mathbb{R} \times \mathbb{R}^+$. Arbitrarily select the reference point $\underline{\theta}_o = [\mu_o, \sigma_o^2] = [0, 1]$ to obtain:

$$\Lambda_{\underline{\theta}}(\underline{x}) = \frac{f_\theta(\underline{x})}{f_{\theta_o}(\underline{x})} = \left(\frac{\sigma_o}{\sigma}\right)^n e^{-n\mu^2/(2\sigma^2)}\ e^{[\mu/\sigma^2,\ -\delta/2]\left[\sum_{i=1}^n x_i,\ \sum_{i=1}^n x_i^2\right]^T},$$

where $\delta = \frac{\sigma_o^2 - \sigma^2}{\sigma^2\sigma_o^2}$. Identifying $\underline{\tau} = \left[\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2\right]$, again condition 1 in Definition 1 is obviously satisfied. Condition 2 in Definition 1 requires a bit more work. While $\Lambda_{\underline{\theta}}(\underline{\tau})$ is no longer an invertible function of τ for for any *single value* of $\underline{\theta} = [\mu, \sigma^2]$, we can find two values $\underline{\theta} \in \{\underline{\theta}_1, \underline{\theta}_2\}$ in Θ for which the vector function $[\Lambda_{\underline{\theta}_1}(\underline{\tau}), \Lambda_{\underline{\theta}_2}(\underline{\tau})]$ of $\underline{\tau}$ is invertible in $\underline{\tau}$. Since this vector is specified by $\Lambda(\underline{x})$, this will imply that $\underline{\tau}$ indexes the likelihood ratios.

To construct this invertible relation denote by $\underline{\lambda} = [\lambda_1, \lambda_2]^T$ an observed pair of samples $[\Lambda_{\underline{\theta}_1}(\underline{\tau}), \Lambda_{\underline{\theta}_2}(\underline{\tau})]^T$ of the surface $\Lambda(\underline{x})$. Now consider the problem of determining $\underline{\tau}$ from the equation $\underline{\lambda} = [\Lambda_{\underline{\theta}_1}(\underline{\tau}), \Lambda_{\underline{\theta}_2}(\underline{\tau})]^T$. Taking the log of both sides and rearranging some terms, we see that this is equivalent to a 2×2 linear system of equations of the form $\underline{\lambda}' = \mathbf{A}\underline{\tau}$, where $\mathbf{A}$ is a matrix involving $\underline{\theta}_o, \underline{\theta}_1, \underline{\theta}_2$ and $\underline{\lambda}'$ is a linear function of $\ln \underline{\lambda}$. You can verify that with the selection of $\underline{\theta}_o = [0, 1]$, $\underline{\theta}_1 = [1, 1]$, $\underline{\theta}_2 = [0, 1/2]$ we obtain $\delta = 0$ or 1 for $\underline{\theta} = \underline{\theta}_1$ or $\underline{\theta}_2$, respectively, and $\mathbf{A} = \mathrm{diag}(1, -1/2)$, an invertible matrix. We therefore conclude that the vector [sample mean, sample variance] indexes the trajectories, and this vector is therefore minimal sufficient.

Example 8 *Minimal sufficient statistic for the location of a Cauchy distribution*

Assume that $X_i \sim f(x;\theta) = \frac{1}{\pi}\frac{1}{1+(x-\theta)^2}$ and, as usual, $\underline{X} = [X_1, \ldots, X_n]^T$ is an i.i.d. sample. Then

$$f(\underline{x};\theta) = \prod_{i=1}^n \frac{1}{\pi}\ \frac{1}{1+(x_i-\theta)^2} = \frac{1}{\pi^n}\ \frac{1}{\prod_{i=1}^n (1+(x_i-\theta)^2)}.$$

Here we encounter a difficulty: the denominator is a $2n$-degree polynomial in θ whose coefficients cannot be determined without specifying the entire set of all possible cross products $x_{i_1}\cdots x_{i_p}$, $p = 1, 2, \ldots, n$, of the x_i's. Since this requires specifying the entire set of sample values there is no finite dimensional sufficient statistic. However, each of these cross products is independent of the ordering of its factors so the ordered statistic $[X_{(1)}, \ldots, X_{(n)}]^T$ is minimally sufficient.

3.5.4 EXPONENTIAL FAMILY OF DISTRIBUTIONS

Let $\underline{\theta} = [\theta_1, \ldots, \theta_p]^T$ take values in some parameter space Θ. The distribution $f_{\underline{\theta}}$ of a random variable X is a member of the p-parameter exponential family if for all $\underline{\theta} \in \Theta$

$$f_{\underline{\theta}}(x) = a(\underline{\theta})b(x)e^{\underline{c}^T(\underline{\theta})\underline{t}(x)}, \quad -\infty < x < \infty \tag{26}$$

for some scalar functions a, b and some p-element vector functions $\underline{c}, \underline{t}$. A similar definition of exponential family holds for vector valued random variables $\underline{X}$, see Bickel and Doksum [7, Ch. 2].

Note that for any $f_{\underline{\theta}}$ in the exponential family its support set $\{x : f_{\underline{\theta}}(x) > 0\}$ does not depend on $\underline{\theta}$. Note that, according to our definition, for $f_{\underline{\theta}}$ to be a member of the p-parameter exponential family the dimension of the vectors $\underline{c}(\underline{\theta})$ and $\underline{t}(x)$ must be exactly p. This is to guarantee that the sufficient statistic has the same dimension as the parameter vector $\underline{\theta}$. While our definition is the most standard [40, 48, 7], some other books, e.g., [55], allow the dimension of the sufficient statistic to be different from p. However, by allowing this we lose some important properties of exponential families [7].

The parameterization of an exponential family of distributions is not unique. In other words, the exponential family is invariant to changes in parameterization. For example, if f_θ, $\theta > 0$, is a member of an exponential family then if one defines $\alpha = 1/\theta$ and $g_\alpha = f_{1/\theta}$ then g_α, $\alpha > 0$, is also in the exponential family, but possibly with different functions $a(\cdot), b(\cdot), c(\cdot)$ and $t(\cdot)$. More generally, if $f_{\underline{\theta}}(\underline{x})$ is a member of the p-dimensional exponential family then transformation of the parameters by any invertible function of $\underline{\theta}$ preserves membership in the exponential family.

To illustrate, let's say that the user redefined the parameters by the mapping $\underline{c} : \theta \longrightarrow \eta$ defined by the invertible transformation $\underline{c}(\underline{\theta}) = \underline{\eta}$. Then, using (26), $f_{\underline{\theta}}$ would be replaced by

$$f_{\underline{\eta}}(x) = \tilde{a}(\underline{\eta}) b(x) e^{\underline{\eta}^T \underline{t}(x)}, \quad -\infty < x < \infty, \tag{27}$$

where $\tilde{a}(\underline{\eta}) = a(\underline{c}^{-1}(\underline{\eta}))$. Thus $f_{\underline{\eta}}$ remains in the exponential family. When expressed in the form (27), the exponential family density $f_{\underline{\eta}}$ is said to be in *canonical form* with *natural parameterization* $\underline{\eta}$. Under the natural parameterization the mean and covariance matrix of the sufficient statistic $\underline{T} = \underline{t}(X)$ are given by (assuming differentiable $\tilde{a}$)

$$E_{\underline{\theta}}[\underline{T}] = \nabla \ln \tilde{a}(\underline{\eta}),$$

and

$$\mathrm{cov}_{\underline{\theta}}[\underline{T}] = \nabla^2 \ln \tilde{a}(\underline{\eta}).$$

For a proof of these relations see Bickel and Doksum [7].

Another parameterization of an exponential family of densities is the *mean value parameterization*. In this parameterization, the functions $\underline{t}(\cdot)$, $a(\cdot)$, $b(\cdot)$ and $\underline{c}(\cdot)$ in (26) are manipulated so that

$$E_{\underline{\theta}}[\underline{T}] = \underline{\theta}. \tag{28}$$

As we will see in the next chapter, when an exponential family is expressed in its mean value parameterization the sufficient statistic $\underline{T}$ is an unbiased minimum variance estimator of $\underline{\theta}$. Thus mean value parameterizations are very special and advantageous.

Examples of distributions in the exponential family include: Gaussian with unknown mean or variance, Poisson with unknown mean, exponential with unknown mean, gamma, Bernoulli with unknown success probability, binomial with unknown success probability, multinomial with unknown cell probabilities. Distributions which *are not* from the exponential family include: Cauchy with unknown median, uniform with unknown support, Fisher-F with unknown degrees of freedom.

When the statistical model is in the exponential family, sufficient statistics for the model parameters have a particularly simple form:

$$f_{\underline{\theta}}(\underline{x}) \quad = \quad \prod_{i=1}^{n} a(\underline{\theta}) b(x_i) e^{\underline{c}^T(\underline{\theta}) \underline{t}(x_i)}$$

$$= \underbrace{a^n(\underline{\theta})\, e^{\underline{c}^T(\underline{\theta}) \overbrace{\sum_{i=1}^n \underline{t}(x_i)}^{\underline{T}}}}_{g(\underline{T},\underline{\theta})} \underbrace{\prod_{i=1}^n b(x_i)}_{h(\underline{x})}$$

Therefore, the following is a p-dimensional sufficient statistic for $\underline{\theta}$

$$\sum_{i=1}^n \underline{t}(X_i) = \left[\sum_{i=1}^n t_1(X_i), \ldots, \sum_{i=1}^n t_p(X_i)\right]^T$$

In fact this is a finite dimensional suff. statistic which is complete and minimal [7].

3.5.5 CHECKING IF A DENSITY IS IN THE EXPONENTIAL FAMILY

Due to the many attractive properties of exponential families, in many situations the first question to be answered is: is the density of my data X a member of this exclusive club? This question might arise, for example, if the input to a known filter or other system has a known density and one can compute a mathematical representation of the density of the output of the filter. To check if the output density is exponential one has to try and manipulate the density into exponential form, as illustrated in the exercises. If this is difficult the next step is to try and show that the density is not in the exponential family. Some properties can be checked immediately, e.g. that the parameters space Θ does not depend on the range of X, e.g. as in a uniform density with unknown region of support boundaries. Another simple test is to compute $\partial^2/\partial\theta\partial x \ln f_\theta(x)$ and verify that it is not of separable form $c'(\theta)t'(x)$ for some functions c and t. This type of question is explored in the exercises.

3.6 BACKGROUND REFERENCES

Mood, Graybill and Boes [48] offers an undergraduate introduction to mathematical statistics with lots of fun exercises and examples. Two of the classic graduate level text books on linear multivariate statistics are Rao [57] and Morrison [50]. Manoukian [43] is a reference book giving a concise compilation of principal results from sampling distribution theory. The book by Johnson *etal* [30], is the first of a set of several volumes of a very comprehensive encyclopedia of probability distributions, random variables, and their properties.

3.7 EXERCISES

3.1 Show that the matrix $\Pi = \mathbf{I}_n - \underline{1}\underline{1}^T/n$ is symmetric and idempotent, where $\mathbf{I}_n$ is the $n \times n$ identity matrix and $\underline{1} = [1, \ldots, 1]^T$ is an n-element column vector of 1's. Show that for $\underline{x} \in \mathbb{R}^n$, $\Pi\underline{x}$ is the vector of residuals $[x_1 - \overline{x_i}, \ldots, x_n - \overline{x_i}]^T$ where $\overline{x_i}$ is the sample mean of elements of $\underline{x}$. Finally show that if $\underline{x}$ has the decomposition $\underline{y} + c\underline{1}$ where $\underline{y}$ has zero (sample) mean and c is an arbitrary scalar, then $\Pi\underline{x} = \underline{y}$, i.e the matrix Π extracts the zero (sample) mean component of $\underline{x}$. It is in this sense that Π is an orthogonal projection matrix onto the space of zero (sample) mean vectors in $\mathbb{R}^n$.

3.2 Assume that a random vector $\underline{X} = [X_1, \ldots, X_n]^T$ has a density $p_{\underline{\theta}}(\underline{x})$ which depends on an unknown parameter vector $\underline{\theta}$. In this exercise you will show that if a statistic $\underline{S} = \underline{S}(\underline{X}) = [S_1(\underline{X}), \ldots, S_k(\underline{X})]^T$ satisfies the Fisher Factorization theorem then the conditional density $p_{\underline{\theta}}(\underline{X}|\underline{S})$ is not a function of $\underline{\theta}$ and thus $\underline{S}$ is a sufficient statistic for $\underline{\theta}$. In the following you should assume that $\underline{X}$ is a discrete random vector and that its joint density $p_{\underline{\theta}}(\underline{x}) = P_{\underline{\theta}}(\underline{X} = \underline{x})$ is a probability mass function (i.e. $p_{\underline{\theta}}(\underline{x}) = 0$ except for a countable number of points $\underline{x} \in \{\underline{x}_1, \underline{x}_2, \ldots\}$ where $p_{\underline{\theta}}(\underline{x}_i) > 0$, and $\sum_{\underline{x}_i} p_{\underline{\theta}}(\underline{x}_i) = 1$).

(a) Use Bayes rule to establish that

$$p_\theta(\underline{x}|\underline{s}) \stackrel{\text{def}}{=} P_{\underline{\theta}}(\underline{X} = \underline{x}|\underline{S} = \underline{s}) = \frac{P_{\underline{\theta}}(\underline{S} = \underline{s}|\underline{X} = \underline{x}) p_{\underline{\theta}}(\underline{x})}{\sum_{\underline{x}_i : \underline{S}(\underline{x}_i) = \underline{s}} p_{\underline{\theta}}(\underline{x}_i)},$$

where the summation of $p_{\underline{\theta}}(\underline{x})$ is over all possible realizations $\{\underline{x}_i\}$ of the vector $\underline{X}$ such that $\underline{S}(\underline{x}_i) = \underline{s}$.

(b) Show that $P_{\underline{\theta}}(\underline{S} = \underline{s}|\underline{X} = \underline{x})$ is equal to one or zero depending on whether $\underline{S}(\underline{x}) = \underline{s}$ or $\underline{S}(\underline{x}) \neq \underline{s}$, respectively. (Hint: express the conditional probability as a ratio and use the definition $\underline{S} = \underline{S}(\underline{X})$ to evaluate the intersection of the events $\underline{S} = \underline{s}$ and $\underline{X} = \underline{x}$).

(c) Using the Fisher Factorization $p_{\underline{\theta}}(\underline{x}) = g_{\underline{\theta}}(\underline{s}) \cdot h(\underline{x})$ show that

$$p_\theta(\underline{x}|\underline{s}) = \begin{cases} \frac{h(\underline{x})}{\sum_{\underline{x}_i : \underline{S}(\underline{x}_i) = \underline{s}} h(\underline{x}_i)}, & \underline{S}(\underline{x}) = \underline{s} \\ 0, & o.w. \end{cases},$$

which, as claimed, does not depend on $\underline{\theta}$.

3.3 Show that the Poisson distribution $p_\lambda(x) = P_\lambda(X = x) = \frac{\lambda^x}{x!} \exp(-\lambda)$, $x = 0, 1, 2, \ldots$ is a member of the one-parameter exponential family. For an i.i.d. sample $\underline{X} = [X_1, \ldots, X_n]^T$ of these Poisson r.v.s find a one dimensional sufficient statistic for λ. Define $\alpha = 1/\lambda$ and show that the reparameterized Poisson distribution $p_\alpha(x)$ is also in the exponential family. Which of these two parameterizations (α or λ) is a mean value paramaterization?

3.4 Let $\underline{X} = [X_1, \ldots, X_n]^T$ be a vector of i.i.d. r.v.s X_i which are uniformly distributed over the interval (θ_1, θ_2), $\theta_1 < \theta_2$. Show that $\underline{S}(\underline{X}) = [\min_i\{X_i\}, \max_i\{X_i\}]^T$ is a sufficient statistic for $\underline{\theta} = [\theta_1, \theta_2]^T$.

3.5 Let Z_i, $i = 1, \ldots, n$, be a set of i.i.d. random variables each with the *alpha density*

$$p_\theta(z) = \frac{\beta}{\sqrt{2\pi}\Phi(\alpha) z^2} \exp\left(-\tfrac{1}{2}\left[\alpha - \beta/z\right]^2\right),$$

where $\beta > 0$ is unknown, α is known and $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$ is the standard normal CDF. The alpha distribution is often used to model tool wear for rotating machinery.

(a) Is the joint density $p_\theta(\underline{z})$ a member of the exponential family of densities?

(b) using the Fisher Factorization find a two dimensional sufficient statistic for estimating the parameter β based on the observation $\underline{Z} = [Z_1, \ldots, Z_n]^T$. Show that this reduces to a one dimensional (scalar) statistic when $\alpha = 0$.

3.6 Let $\underline{X} = [X_1, \ldots, X_n]^T$ be a vector of i.i.d. Gaussian r.v.s with mean μ and variance $\sigma^2 = \mu^2$ ($X_i \sim \mathcal{N}(\mu, \mu^2)$).

(a) Show that the sample mean $\overline{X}_i = \frac{1}{n}\sum_{i=1}^n X_i$ is *not* a sufficient statistic for μ by demonstrating that the conditional jpdf of $\underline{X}$ given $\overline{X}$ is a function of μ.

(b) Find a two dimensional sufficient statistic.

3.7 Let $T = T(\underline{x})$ be a sufficient statistic for θ, where $\underline{x} \sim f(\underline{x};\theta) = g(T(\underline{x}),\theta)h(\underline{x})$ is a discrete random variable. Show that T has probability mass function

$$f(t;\theta) = g(t,\theta)q(t),$$

where

$$q(t) = \sum_{\{\underline{x}:T(\underline{x})=t\}} h(\underline{x}).$$

3.8 Consider the case that $\underline{X} = [X_1,\ldots,X_n]^T$ are drawn from a Bernoulli distribution, $X_i \in \{0,1\}$, $P(X_i = 1) = 1 - P(X_i = 0) = p$, $p \in [0,1]$, and X_i's are i.i.d. Show that the Binomial r.v. $T = \sum_{i=1}^n X_i$ is a sufficient statistic for p. Show that T is minimal. Also show that T is a complete sufficient statistic (Hint: for any function g express $E_\theta[g(T)]$ as a polynomial in $\theta = p$ and compute n-th order derivative wrt p).

3.9 Let $X_1,\ldots,X_n$ be i.i.d. uniform r.v.s having common density $f_{X_i}(x;\theta) = \frac{1}{\theta}I_{[0,\theta]}(x)$ $(\theta > 0)$, where $I_A(x)$ denotes the indicator function of the set A. Show that $T = \max(X_1,\ldots,X_n)$ is a complete sufficient statistic for θ by the following steps:

(a) Show the sufficiency of T.

(b) Derive the density function of T.

(c) Show that $E_\theta[g(T)] = 0$, for all $\theta > 0$ implies g is identically zero.

End of chapter

4 FUNDAMENTALS OF PARAMETRIC ESTIMATION

In the last chapter we explored the foundation of statistical inference: the formulation of a statistical model and sufficient statistics for model parameters. In this chapter we go on to develop explicit methods to estimate the parameters from random samples from the model, paying close attention to how well the accuracy of these estimates hold up over different sample realizations.

We will start off with the basic mathematical formulation of estimation and then, specializing to the case of scalar one-dimensional parameters, consider two different models: random parameters and non-random parameters. It turns out, perhaps surprisingly, that estimation of random parameters has a cleaner theory. This is because for random parameters one can more straightforwardly assess the estimator's mean accuracy and specify procedures for finding optimal estimators, called Bayes estimators, having highest possible accuracy. In particular we define three different optimality criteria mean squared error (MSE), mean absolute error (MAE), and mean uniform error, also called probability of large error (P_e). We then turn to deterministic scalar parameters for which we focus on bias and variance as measures of estimator accuracy. This leads to the concept of Fisher information and the Cramèr-Rao lower bound on variance of unbiased estimators. Finally we generalize the treatment to multiple (vector) parameters.

4.1 ESTIMATION: MAIN INGREDIENTS

We follow the same notation as in the last chapter, summarized below.

$\underline{X} \in \mathcal{X}$ is a random measurement or observation
$\mathcal{X}$ is the sample space of measurement realizations $\underline{x}$
$\underline{\theta} \in \Theta$ is an unknown parameter vector of interest
$\Theta \subset \mathbb{R}^p$ is the parameter space
$f(\underline{x};\underline{\theta})$ is the pdf of $\underline{X}$ for given $\underline{\theta}$ (a known function)

With these definitions, the objective of parameter estimation is to design an estimator function

$$\hat{\underline{\theta}} = \hat{\underline{\theta}}(x)$$

which maps $\mathcal{X}$ to $\mathbb{R}^p \supset \Theta$. The concept is illustrated in Fig. 1.

It is important to distinguish between an estimator, which is a function of the sample $\underline{X}$, and an estimate, which is an evaluation of the function at a particular realization $\underline{x}$ of $\underline{X}$, i.e.:

- the function $\hat{\underline{\theta}}$ is an *estimator*.
- the point $\hat{\underline{\theta}}(x)$ is an *estimate*.

A natural question arises. What is an appropriate design criterion for constructing an estimator? There are many possible approaches to this. In this chapter we will describe two of the principal approaches. The first assumes that $\underline{\theta}$ is random and the second assumes it is deterministic. Common to both approaches is the specification of a loss function, also called a *risk* function, associated with an estimator that measures the estimation error as a function of both the sample and the parameter values.

Define $c(\hat{\underline{\theta}}(x);\underline{\theta})$ a loss function associated with $\hat{\underline{\theta}}$ for given $\underline{\theta}$ and $\underline{X} = \underline{x}$. The optimum estimator, should it exist, might be found by minimizing average loss $E[C]$, where as usual, the capitalization C denotes the random variable $c(\hat{\underline{\theta}}(\underline{X}), \underline{\theta})$.

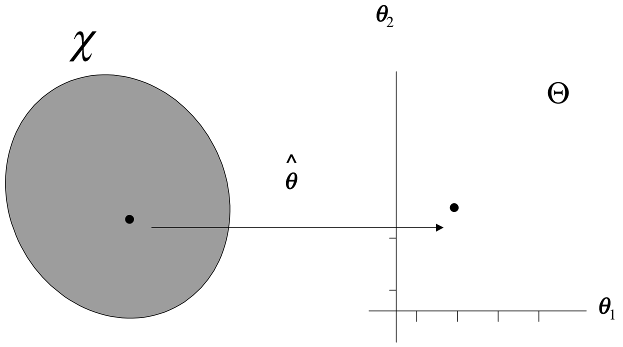


Figure 1: *An estimator of a p-dimensional parameter $\underline{\theta}$ given an n-dimensional random sample $\underline{X}$ is a mapping of $\mathcal{X}$ to $\mathbb{R}^p$*

4.2 ESTIMATION OF RANDOM SCALAR PARAMETERS

For the case that $\underline{\theta}$ is a random scalar parameter θ we have access to the following information:

$f(\theta)$: a prior p.d.f. for $\underline{\theta}$.

$f(\underline{x}|\theta)$: a conditional p.d.f.

$f(\theta|\underline{x})$: the posterior p.d.f. for θ that is determined by Bayes rule:

$$f(\theta|\underline{x}) = \frac{f(\underline{x}|\theta)f(\theta)}{f(\underline{x})}.$$

$f(\underline{x})$: the marginal p.d.f. determined by marginalization over θ

$$f(\underline{x}) = \int_\Theta f(\underline{x}|\underline{\theta})f(\underline{\theta})d\underline{\theta}$$

With the above we can compute the average loss, also called Bayes risk, as

$$E[C] = \int_\Theta \int_{\mathcal{X}} c(\hat{\underline{\theta}}(\underline{x}), \underline{\theta}) f(\underline{x}|\underline{\theta}) f(\underline{\theta}) \; d\underline{x} d\underline{\theta}.$$

We now can naturally define an optimal estimator. A scalar estimator $\hat{\theta}$ which minimizes the average loss is called a *Bayes estimator*. Some reasonable loss functions for this estimation problem are

$c(\hat{\theta};\theta) = |\hat{\theta} - \theta|^2$: squared error

$c(\hat{\theta};\theta) = |\hat{\theta} - \theta|$: absolute error

$c(\hat{\theta};\theta) = I(|\hat{\theta} - \theta| > \epsilon)$: uniform error

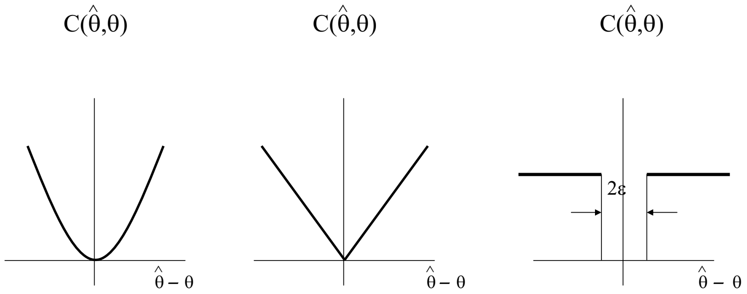


Figure 2: *Three loss functions for scalar parameter estimation: (a) squared error, (b) absolute error, (c) uniform error.*

Figure 2 illustrates these three loss functions as a function of the estimator error difference $\hat{\theta} - \theta$.

For each of the three loss functions we can compute the mean loss and obtain the Bayes risk functions (functions of $f(\theta), f(\underline{x}|\theta)$ and $\hat{\theta}$):

Estimator MSE:

$$\text{MSE}(\hat{\theta}) = E[|\hat{\theta} - \theta|^2]$$

Estimator MAE:

$$\text{MAE}(\hat{\theta}) = E[|\hat{\theta} - \theta|]$$

Error Probability:

$$P_e(\hat{\theta}) = P(|\hat{\theta} - \theta| > \epsilon)$$

It remains to find the estimators $\hat{\theta}$, called *optimal estimators*, which minimize each of these criteria.

4.2.1 MINIMUM MEAN SQUARED ERROR ESTIMATION

The MSE is the most widespread estimation criterion and arguably the one with the longest history. The optimal minimum mean squared error estimator (MMSEE) is the *conditional mean estimator* (CME) defined as

$$\hat{\theta}(\underline{X}) = E[\theta|\underline{X}] = \text{mean}_{\theta\in\Theta}\{f(\theta|\underline{X})\},$$

where

$$\text{mean}_{\theta\in\Theta}\{f(\theta|\underline{X})\} = \int_{-\infty}^{\infty} \theta f(\theta|\underline{X}) d\theta.$$

The CME has an intuitive mechanical interpretation as the center of mass (1st moment of inertia) of the mass density $f(\theta|\underline{x})$ (Fig. 3). The CME corresponds to the posterior average value of the parameter after you have observed the data sample.

The CME satisfies an orthogonality condition: the Bayes estimator error is orthogonal to any (linear or non-linear) function of the data. This condition is mathematically expressed below for the general case of complex rv's,

$$E[(\theta - \hat{\theta}(\underline{X}))g(\underline{X})^*] = 0,$$

for any function g of x. Here u^* denotes complex conjugate of u.

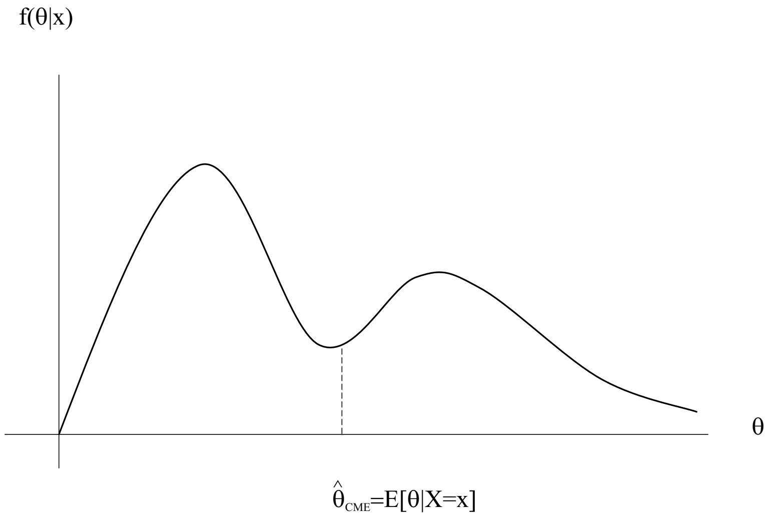


Figure 3: *Conditional mean estimator minimizes MSE*

Proof: Write the MSE as

$$E[|\hat{\theta} - \theta|^2] = E[|(\hat{\theta} - E[\theta|\underline{X}]) - (\theta - E[\theta|\underline{X}])|^2]$$

$$= E[|\hat{\theta} - E[\theta|\underline{X}]|^2] + E[|\theta - E[\theta|\underline{X}]|^2]$$

$$-E[g(\underline{X})^*(\theta - E[\theta|\underline{X}])] - E[g(\underline{X})(\theta - E[\theta|\underline{X}])^*]$$

where $g(\underline{X}) = \hat{\theta} - E[\theta|\underline{X}]$ is a function of $\underline{X}$ only.

Step 1: show orthogonality condition

$$\begin{aligned} E[g(\underline{X})(\theta - E[\theta|\underline{X}])] &= E\left[\ E[g(\underline{X})(\theta - E[\theta|\underline{X}])^* |\ \underline{X}]\ \right] \\ &= E\left[g(\underline{X})\ \underbrace{E\left[\theta - E[\theta|\underline{X}]\ |\ \underline{X}\right]}_{=0} \right] = 0 \end{aligned}$$

Step 2: Next show $E[\theta|\underline{X}]$ minimizes MSE

$$
\begin{aligned}
E[|\hat{\theta}-\theta|^2] &= E[|\hat{\theta}-E[\theta|\underline{X}]|^2] + E[|\theta - E[\theta|\underline{X}]|^2] \\
&\geq E[|\theta - E[\theta|\underline{X}]|^2]
\end{aligned}
$$

where "=" occurs iff $\hat{\theta} = E[\theta|\underline{X}]$ ◇

4.2.2 MINIMUM MEAN ABSOLUTE ERROR ESTIMATOR

For convenience we assume θ is a real valued scalar and $F(\theta|\underline{x}) = \int^{\theta} f(\theta'|\underline{x})d\theta'$ is a continuous function of θ. The minimal mean absolute error estimator (MMAEE) is the conditional median estimator (CmE)

$$\hat{\theta}(\underline{X}) = \text{median}_{\theta\in\Theta}\{f(\theta|\underline{X})\},$$

where

$$
\begin{aligned}
\text{median}_{\theta\in\Theta}\{f(\theta|\underline{X})\} &= \min\{u : \int_{-\infty}^{u} f(\theta|\underline{X})d\theta = 1/2\} \qquad (29)\\
&= \min\left\{u : \int_{-\infty}^{u} f(\underline{X}|\theta)f(\theta)d\theta = \int_{u}^{\infty} f(\underline{X}|\theta)f(\theta)d\theta\right\}. \qquad (30)
\end{aligned}
$$

The median of a density separates the density into two halves of equal mass (Fig. 4). When $F(\theta|\underline{x})$ is strictly increasing over Θ the "min" in the definition of the median is not necessary - but it may be required when there are regions of Θ where the density $f(\theta|\underline{x})$ is equal to zero. If $f(\theta|\underline{X})$ is continuous in θ the CmE also satisfies an orthogonality condition:

$$E[\text{sgn}(\theta - \hat{\theta}(\underline{X}))g(\underline{X})] = 0,$$

and thus for minimum MAE estimation it is the sign of the optimum estimation error that is orthogonal to any function of the data sample.

Proof: Let $\hat{\theta}_m$ = median of $f(\theta|\underline{X})$.

Then by definition of median for continuous densities

$$
\begin{aligned}
E[\text{sgn}(\theta-\hat{\theta}_m) \mid \underline{X}] &= \int_{\Theta} \text{sgn}(\theta - \hat{\theta}_m(\underline{X}))\ f(\theta|\underline{X})d\theta \\
&= \int_{\theta>\hat{\theta}_m(\underline{X})} f(\theta|\underline{X})d\theta - \int_{\theta\leq\hat{\theta}_m(\underline{X})} f(\theta|\underline{X})d\theta \\
&= 0
\end{aligned}
$$

Step 1: show orthogonality condition:

$$E[\text{sgn}(\theta-\hat{\theta}_m)g(\underline{X})] = E[\ \underbrace{E[\text{sgn}(\theta-\hat{\theta}_m)|\underline{X}]}_{=0}\ g(\underline{X})]$$

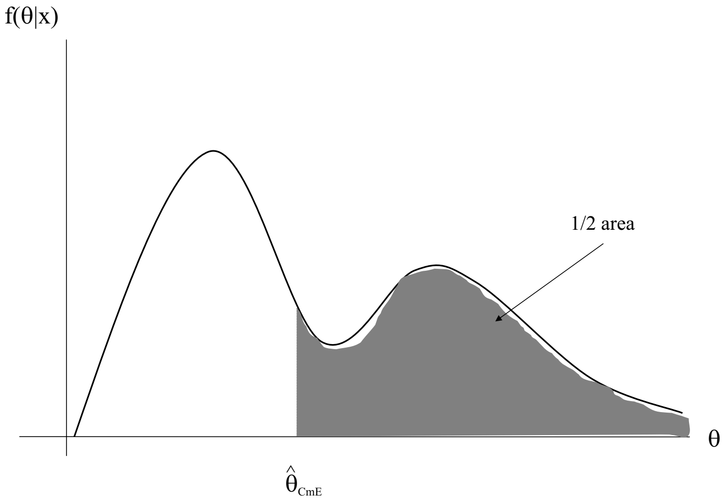


Figure 4: *Conditional median estimator minimizes MAE*

Step 2: for $\hat{\theta}$ arbitrary we have (apply "useful formula" below)

$$\begin{aligned}
\mathrm{MAE}(\hat{\theta}) &= E[|\underbrace{\theta-\hat{\theta}_m}_{a}+\underbrace{\hat{\theta}_m-\hat{\theta}}_{\Delta}|] \\
&= E[|\theta-\hat{\theta}_m|] + \underbrace{E[\mathrm{sgn}(\theta-\hat{\theta})\Delta]}_{=0} \\
&\quad +E\underbrace{[\mathrm{sgn}(a+\Delta)-\mathrm{sgn}(a)](a+\Delta)}_{\geq[\mathrm{sgn}(a+\Delta)-1](a+\Delta)\geq 0} \\
&\geq E[|\theta-\hat{\theta}_m|]
\end{aligned}$$

Useful formula: $|a+\Delta| = |a| + \mathrm{sgn}(a)\Delta + [\mathrm{sgn}(a+\Delta) - \mathrm{sgn}(a)](a+\Delta)$

4.2.3 MINIMUM MEAN UNIFORM ERROR ESTIMATION

Unlike the MSE or MAE, the MUE penalizes only those errors that exceed a tolerance level $\epsilon > 0$ and this penalty is uniform. For small ϵ the optimal estimator is the *maximum a posteriori* (MAP) estimator, which is also called the *posterior mode* estimator (Fig. 5)

$$\begin{aligned}
\hat{\theta}(\underline{X}) &= \mathrm{argmax}_{\theta\in\Theta}\{f(\theta|\underline{X})\} && (31)\\
&= \mathrm{argmax}_{\theta\in\Theta}\left\{\frac{f(\underline{X}|\theta)f(\theta)}{f(\underline{X})}\right\} && (32)\\
&= \mathrm{argmax}_{\theta\in\Theta}\{f(\underline{X}|\theta)f(\theta)\}. && (33)
\end{aligned}$$

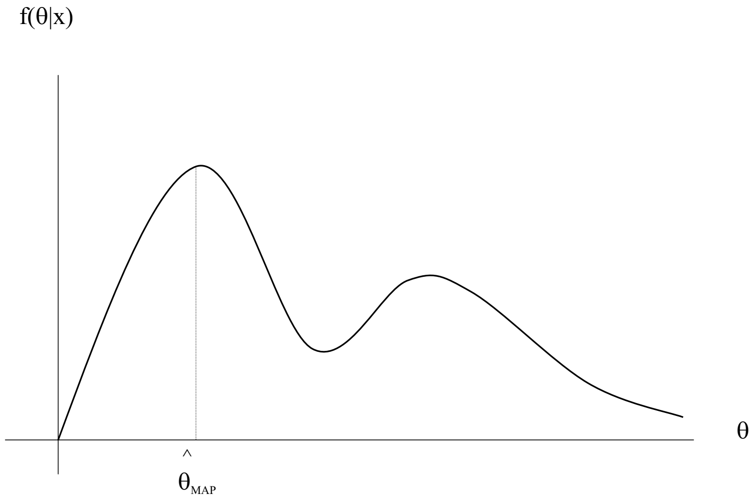


Figure 5: *Maximum a posteriori estimator minimizes* P_e

Notice that the third line of (33) is best suited to computation of the MAP estimator since it does not require the marginal $f(\underline{x})$, which can be difficult to compute.

Proof:

Assume that ϵ is a small and positive number. The probability that the magnitude estimator error exceeds ϵ is simply expressed

$$\begin{aligned} P_e(\hat{\theta}) &= 1 - P(|\theta - \hat{\theta}| \leq \epsilon) \\ &= 1 - \int_{\mathcal{X}} d\underline{x} f(\underline{x}) \int_{\{\theta: |\theta - \hat{\theta}(\underline{x})| \leq \epsilon\}} f(\theta|\underline{x}) d\theta. \end{aligned}$$

Consider the inner integral (over θ) in the above expression. This is an integral over θ within a window, which we call the *length* 2ϵ *window*, centered at $\hat{\theta}$. Referring to Fig. 6, it should be evident to the reader that, if ϵ is sufficiently small, this integral will be maximized by centering the length 2ϵ window at the value of θ that maximizes the integrand $f(\theta|\underline{x})$. This value is of course the definition of the MAP estimate $\hat{\theta}$. ◇

Now that we have seen three different estimator criteria, and their associated optimal estimators, we make several general remarks.

1. The CmE may not exist for discrete Θ since the median may not be well defined.
2. Only the CME requires (often difficult) computation of the normalization factor $f(\underline{x})$ in the posterior $f(\theta|\underline{x}) = f(\underline{x}|\theta)/f(\underline{x})$.
3. Each of these estimators depends on x only through posterior $f(\theta|\underline{x})$.
4. When the posterior is continuous, unimodal, and symmetric then each of the above estimators are identical (VanTrees [73])! See Fig. 7 for illustration.
5. If $T = T(\underline{X})$ is a sufficient statistic the posterior depends on $\underline{X}$ only through T. Indeed, if $f(\underline{X}|\theta) = g(T;\theta)h(\underline{X})$, then by Bayes rule

$$f(\theta|\underline{X}) = \frac{f(\underline{X}|\theta)f(\theta)}{\int_{\Theta} f(\underline{X}|\theta)f(\theta)d\theta} = \frac{g(T;\theta)f(\theta)}{\int_{\Theta} g(T;\theta)f(\theta)d\theta}$$

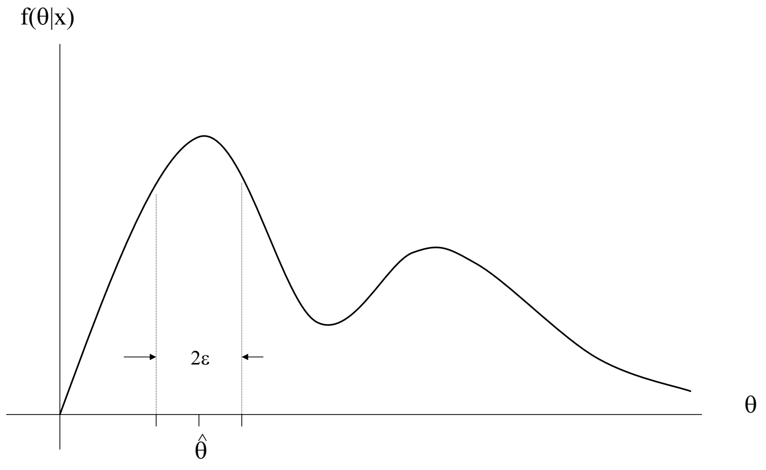


Figure 6: *Posterior density integrated over window of length* 2ϵ

which is only a function of $\underline{X}$ through T. Thus, in terms of optimal estimation performance, one loses nothing by compressing $\underline{X}$ to a sufficient statistic.

6. The CME has the following linearity property. For any random parameter variables θ_1 and θ_2: $E[\theta_1 + \theta_2|\underline{X}] = E[\theta_1|\underline{X}] + E[\theta_2|\underline{X}]$. This property is not shared by the CmE or the MAP estimator.

4.2.4 BAYES ESTIMATOR EXAMPLES

Here we give four examples of statistical models, priors, and derive their optimal estimators under various criteria.

These are the examples we will cover (hotlinks on the web version)

* Estimation of width of uniform density

* Estimation of a Gaussian signal

* Estimation of magnitude of Gaussian signal

* Estimation of a binary signal in Gaussian noise

Example 9 *ESTIMATION OF WIDTH OF UNIFORM PDF*

Consider the following motivating problem. A networked computer terminal takes a random amount of time to connect to another terminal after sending a connection request at time $t = 0$. You, the user, wish to schedule a transaction with a potential client as soon as possible after sending the request. However, if your machine does not connect within the scheduled time then your client will go elsewhere. If one assumes that the connection delay is a random variable X that is uniformly distributed over the time interval $[0, \theta]$ you can ensure your client that the delay will not exceed θ. The problem is that you do not know θ so it must be estimated from past

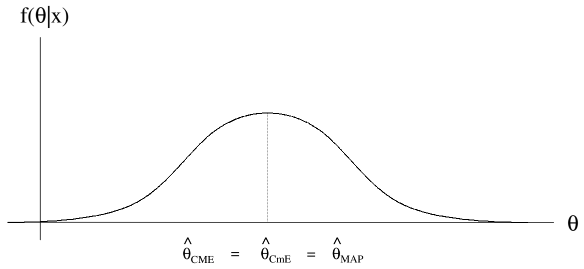


Figure 7: *Symmetric and continuous posterior density*

experience, e.g., the sequence of previously observed connection delays $X_1, \ldots, X_n$. By assuming a prior distribution on θ an optimal estimate can be obtained using the theory developed above.

So now let's formulate this in our language of estimation theory.

We assume that $X_1, \ldots, X_n$ are conditionally i.i.d. uniform samples each with conditional density

$$f(x_1|\theta) = \frac{1}{\theta} I_{[0,\theta]}(x_1).$$

Let's say that based on your experience with lots of different clients you determine that a reasonable prior on θ is

$$f(\theta) = \theta\, e^{-\theta}, \quad \theta > 0.$$

Figure 8 illustrates these two densities.

We will derive the CME, CmE, and MAP estimators of θ. There are two steps.

Step 1: Find the posterior $f(\theta|\underline{x}) = f(\underline{x}|\theta)f(\theta)/f(\underline{x})$

$$\begin{aligned}
f(\underline{x}|\theta)f(\theta) &= \left(\prod_{i=1}^n \frac{1}{\theta} I_{[x_i,\infty)}(\theta)\right)\left(\theta e^{-\theta}\right) \\
&= \frac{e^{-\theta}}{\theta^{n-1}} \underbrace{\prod_{i=1}^n I_{[x_i,\infty)}(\theta)}_{I_{[x_{(1)},\infty)}(\theta)} \\
&= \frac{e^{-\theta}}{\theta^{n-1}} I_{[x_{(1)},\infty)}(\theta).
\end{aligned}$$

where $x_{(1)} = \max\{x_i\}$. Observe that the function $\frac{e^{-\theta}}{\theta^{n-1}}$ is monotone decreasing over $\theta > 0$ (verify that the derivative of its logarithm is negative).

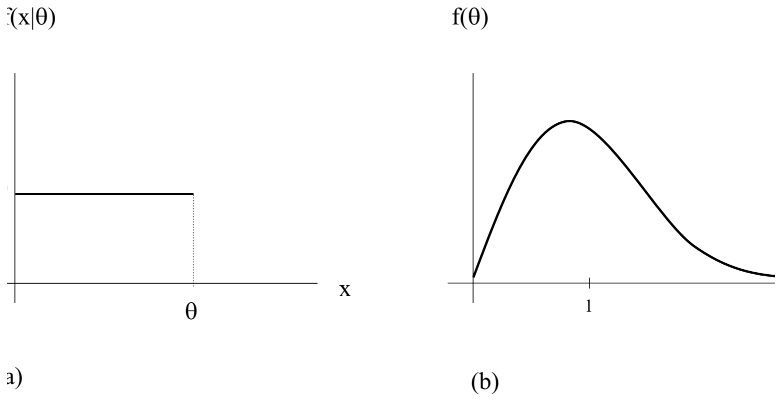


Figure 8: *(a) Uniform density of unknown width θ, (b) prior on θ*

Furthermore,

$$\begin{aligned} f(\underline{x}) &= \int_0^\infty f(\underline{x}|\theta)f(\theta)d\theta \\ &= q_{-n+1}(x_{(1)}) \end{aligned}$$

where q_n is the monotone decreasing function

$$q_n(x) \stackrel{\text{def}}{=} \int_x^\infty \theta^n e^{-\theta} d\theta$$

Recursive formula: $q_{-n-1}(x) = \frac{1}{n}\left(\frac{1}{x^n}e^{-x} - q_{-n}(x)\right)$, $n = 0, -1, -2, \ldots$.
Step 2: find optimal estimator functions:

$$\begin{aligned} \hat{\theta}_{\text{MAP}} &= X_{(1)} \\ \hat{\theta}_{\text{CME}} &= q_{-n+2}(X_{(1)})/q_{-n+1}(X_{(1)}) \\ \hat{\theta}_{\text{CmE}} &= q_{-n+1}^{-1}\left(\tfrac{1}{2}q_{-n+1}(X_{(1)})\right). \end{aligned}$$

Note that only the MAP estimator is a simple function of $\underline{X}$ while the two others require more difficult computation of integrals q_n and/or an inverse function q_n^{-1}. These estimators are illustrated in Fig. 9 along with the posterior density $f(\theta|\underline{x})$.

Example 10 *ESTIMATION OF GAUSSIAN AMPLITUDE*

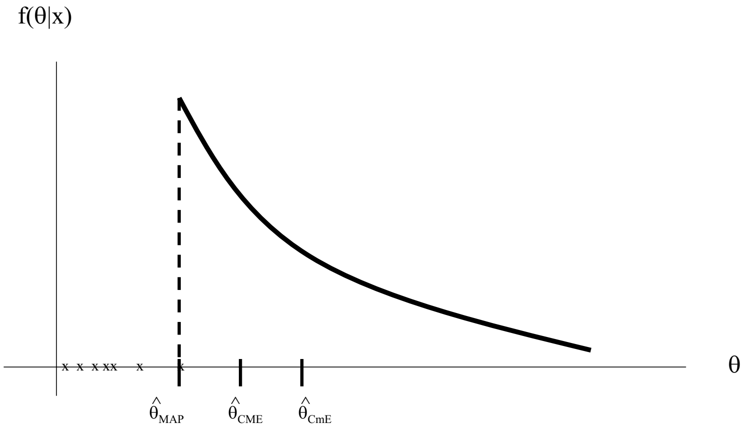


Figure 9: *The estimators CME, CmE and MAP for the width parameter θ of the underlying uniform density with prior given by Fig. 8.b.*

A very common assumption arising in many signal extraction problems is the assumption of a Gaussian distributed signal observed in additive Gaussian noise. For example, a radar target acquisition system might transmit a pulse to probe for possible targets in a cell located at a particular point in space. If a strong reflecting target is present at that point then it reflects some of the energy in the radar pulse back to the radar, resulting in a high energy signal, called a radar return, at the radar receiver. The amplitude of this signal might contain useful information about the identity of the target. Estimation of the radar return is complicated by the presence of ambient noise generated in the radar receiver (thermal noise) or by interference from other sources (clutter) in the cell. Based on field trials of the radar system prior mean and variances of the received signal and the noise might be available.

To set this up more formally as an estimation problem we define two jointly Gaussian r.v.s: S, X with known means, variances, and covariance

$$\begin{aligned} &E[S] = \mu_S, \ E[X] = \mu_X, \\ &\mathrm{var}(S) = \sigma_S^2, \ \mathrm{var}(X) = \sigma_X^2 \\ &\mathrm{cov}(S, X) = \rho \, \sigma_S \sigma_X. \end{aligned}$$

S will play the role of the signal and X will be the measurement. Of course the specific form of the covariance function will depend on the receiver structure, e.g., it reduces to a simple function of σ_S and σ_X for an additive noise model.

The objective is to find an optimal estimator of S given measured X. As in the previous example the derivation of CME, CmE and MAP estimators is divided into two parts.

Step 1: find the posterior density.

A fundamental fact about jointly Gaussian random variables is that if you condition on one of the variables then the other variable is also Gaussian, but with different mean and variance equal to its conditional mean and variance (see Fig. 11 and Exercise 4.25 at the end of chapter). In particular, the conditional density of S given $X = x$ is Gaussian with mean parameter

$$\mu_{S|X}(x) = E[S|X = x] \quad = \quad \mu_S + \rho \frac{\sigma_S}{\sigma_X}(x - \mu_X),$$

and variance parameter

$$\sigma^2_{S|X} = E[(S - E[S|X])^2|X = x] \quad = \quad (1 - \rho^2)\sigma^2_S,$$

so that the conditional density takes the form

$$\begin{aligned} f_{S|X}(s|x) &= \frac{f_{X|S}(x|s) f_S(s)}{f_X(x)} \\ &= \frac{1}{\sqrt{2\pi\sigma^2_{S|X}}} \exp\left\{ -\frac{\left(s - \mu_{S|X}(x)\right)^2}{2\sigma^2_{S|X}} \right\}. \end{aligned}$$

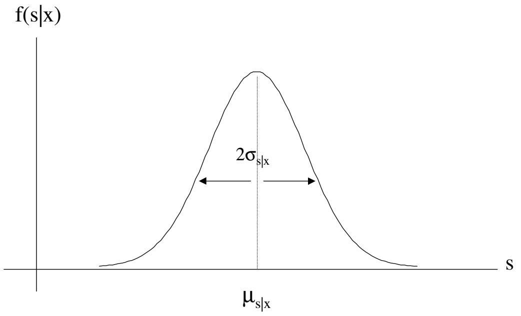


Figure 10: *The posterior $f(s|x)$ when s, x are jointly Gaussian is a Gaussian density.*

Step 2: find the form of the optimal estimators

We immediately note that, as the posterior is continuous, symmetric and unimodal, the MAP, CME, and CmE estimators are of identical form. Bringing out the explicit dependency of the estimator $\hat{S}$ on the observed realization x we have:

$$\hat{S}(x) = \mu_{S|X}(x) = \text{ linear in } x.$$

An interesting special case, relevant to the radar example discussed above, is the independent additive noise model where $X = S + V$. For this case $\sigma^2_X = \sigma^2_S + \sigma^2_V$, $\rho^2 = \sigma^2_S/(\sigma^2_S + \sigma^2_V)$ and therefore

$$\hat{S}(x) \quad = \quad \mu_S + \frac{\sigma^2_S}{\sigma^2_S + \sigma^2_V}(x - \mu_X).$$

We can easily derive the performance of the optimal estimator under the MSE criterion

Minimum MSE: $E[(S-\hat{S})^2] = (1-\rho^2)\sigma_S^2$.

A little more work produces expressions for the performances of this optimal estimator under the MAE and Pe (MUE) criteria:

Minimum MAE: $E[|S-\hat{S}|] = \sqrt{(1-\rho^2)\sigma_S^2}\,\sqrt{\frac{2}{\pi}}$

Minimum Pe: $P(|S-\hat{S}| > \epsilon) = 1 - \text{erf}\left(\epsilon/\sqrt{2(1-\rho^2)\sigma_S^2}\right)$

Example 11 *Estimation of magnitude of Gaussian signal*

Now we change Example 10 a little bit. What if the radar operator was only interested in the energy of the received signal and not its sign (phase)? Then the proper objective would be to estimate the magnitude $|S|$ instead of the magnitude and phase S. Of course, an ad hoc estimation procedure would be to simply take the previously derived estimator $\hat{S}$ and use its magnitude $|\hat{S}|$ to estimate $|S|$ but is this the best we can do?

Let's see what the form of the best estimators of $|S|$ are.

Again we define two jointly Gaussian r.v.s: S, X with means, variances, and covariance

$$\begin{aligned} E[S] &= \mu_S, \; E[X] = \mu_X, \\ \text{var}(S) &= \sigma_S^2, \; \text{var}(X) = \sigma_X^2, \\ \text{cov}(S,X) &= \rho\;\sigma_S\sigma_X. \end{aligned}$$

Now the objective is to estimate the random variable $Y = |S|$ based on X. Note: the pair Y, X no longer obeys a jointly Gaussian model. But, using first principles, we can easily derive the optimal estimators. The first step is to compute the posterior density $f_{Y|X}$.

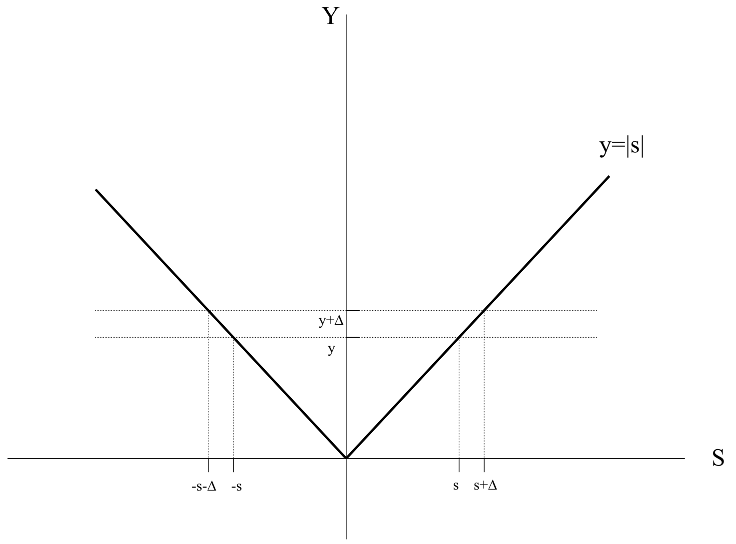


Figure 11: *Illustration of the method of differentials for finding conditional density of $Y = |S|$ given X from the probability $P(y < Y \leq y + \Delta|X = x) \approx f_{Y|X}(y|x)\Delta$, $0 < \Delta \ll 1$.*

Since we know $f_{S|X}$ from the previous example this is a simple transformation of variables problem of elementary probability. We use the method of differentials (see Fig. 11) to obtain the following relation, valid for small Δ

$$f_{Y|X}(y|x)\Delta = f_{S|X}(y|x)\Delta + f_{S|X}(-y|x)\Delta, \quad y \geq 0,$$

or more explicitly

$$\begin{aligned} f_{Y|X}(y|x) &= \\ & \frac{1}{\sqrt{2\pi\sigma^2_{S|X}}}\left(\exp\left\{-\frac{\left(y-\mu_{S|X}(x)\right)^2}{2\sigma^2_{S|X}}\right\} + \exp\left\{-\frac{\left(y+\mu_{S|X}(x)\right)^2}{2\sigma^2_{S|X}}\right\}\right) I_{[0,\infty)}(y). \quad (34)\end{aligned}$$

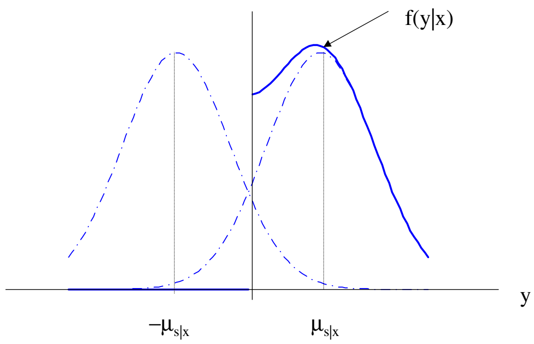


Figure 12: *Posterior density of $Y = |S|$ given X*

Unlike Example 10 this posterior density, shown in Fig. 12 is no longer symmetric in y. Hence we expect the CME, CmE, and MAP estimators to be different.

The CME can be derived in explicit closed form by integration over $y \in [0, \infty)$ of the function $yf_{Y|X}(y|x)$ specified in (34)

$$\hat{Y}_{\mathrm{CME}}(x) = E[Y|X=x] \quad = \quad \left|\mu_{S|X}(x)\right| \ \mathrm{erf}\left(\frac{|\mu_{S|X}(x)|}{\sigma_{S|X}\sqrt{2}}\right) + \sqrt{\frac{2}{\pi}}\sigma_{S|X}\, e^{-\mu_S^2/2\sigma^2_{S|X}}.$$

On the other hand, by investigating the MMAE equation $\int_{\hat{Y}}^{\infty} f_{Y|X}(y|x)dy = \int_0^{\hat{Y}} f_{Y|X}(y|x)dy$ it is easily seen that the CmE can only be implicitly given as the solution $\hat{Y} = \hat{Y}_{\mathrm{CmE}}$ of the following

$$\mathrm{erf}\left(\frac{\hat{Y}-\mu_{S|X}(x)}{\sigma_{S|X}\sqrt{2}}\right) + \mathrm{erf}\left(\frac{\hat{Y}+\mu_{S|X}(x)}{\sigma_{S|X}\sqrt{2}}\right) = \frac{1}{2}.$$

Finally, as $f_{Y|X}(y|x)$ is concave and smooth in y, the MAP estimator $\hat{Y} = \hat{Y}_{\text{MAP}}$ occurs at a stationary point in y of the so called "MAP equation"

$$0 = \frac{\partial f(y|x)}{\partial y}.$$

Using (34) this yields

$$\hat{Y}(x) = \mu_{S|X}(x) \frac{\exp\left\{-\frac{\left(\hat{Y}-\mu_{S|X}(x)\right)^2}{2\sigma^2_{S|X}}\right\} - \exp\left\{-\frac{\left(\hat{Y}+\mu_{S|X}(x)\right)^2}{2\sigma^2_{S|X}}\right\}}{\exp\left\{-\frac{\left(\hat{Y}-\mu_{S|X}(x)\right)^2}{2\sigma^2_{S|X}}\right\} + \exp\left\{-\frac{\left(\hat{Y}+\mu_{S|X}(x)\right)^2}{2\sigma^2_{S|X}}\right\}}.$$

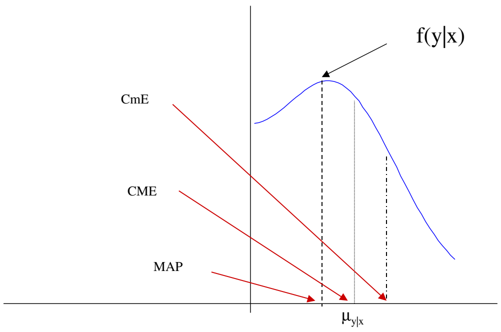


Figure 13: *Three optimal estimators of $Y = |S|$ when S, X are jointly Gaussian.*

The above optimal estimators are illustrated in Fig. 13. It can be verified that as $\mu_{S|X}/\sigma_{S|X} \to \infty$ all three estimators converge to an identical limit:

$$\hat{Y}(x) \to \left|\mu_{S|X}(x)\right|.$$

This limiting case occurs since the posterior density becomes a dirac delta function concentrated at $y = \mu_{S|Y}(x)$ as $\mu_{S|X}/\sigma_{S|X} \to \infty$. Observe that none of these estimators of $|S|$ are given by $|\hat{S}|$ where $\hat{S}$ is the corresponding MAP/CME/CmE estimate of S derived in Example 10. This illustrates an important fact: estimation of random parameters is not invariant to functional transformation,

Example 12 *Estimation of sign of Gaussian signal*

Above we derived optimal estimators for magnitude of a Gaussian random variable based on Gaussian observations. Well, how about when only the phase is of interest, e.g., when the radar operator wants to estimate the sign as opposed to the magnitude of the signal? We treat a simplified version of this problem in this example.

Assume that the model for the observation is

$$X = \theta + W$$

where W is a zero mean Gaussian noise with variance σ^2 and θ is an equally likely binary random variable: $P(\theta = 1) = P(\theta = -1) = \frac{1}{2}$, $\Theta = \{-1, 1\}$. This corresponds to our radar problem when the prior mean μ_S is zero (why?) and an additive noise model is assumed.

Here the posterior density is a probability mass function since the signal θ is discrete valued:

$$p(\theta|x) = \frac{f(x|\theta)p(\theta)}{f(x)},$$

where $p(\theta) = 1/2$. For convenience we have eliminated subscripts on densities. Furthermore, as illustrated in Fig. 14,

$$f(x|\theta) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{(x-1)^2}{2\sigma^2}\right), & \theta = 1 \\ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{(x+1)^2}{2\sigma^2}\right), & \theta = -1 \end{cases}.$$

Hence

$$f(x) = f(x|\theta = 1)\tfrac{1}{2} + f(x|\theta = -1)\tfrac{1}{2}.$$

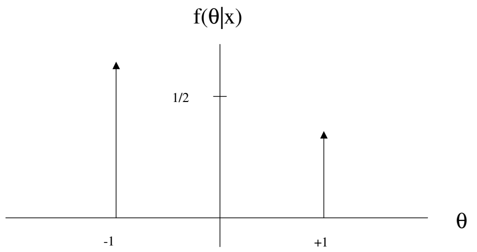


Figure 14: *The posterior density $f(\theta|x)$ concentrates mass on the pair of points $\theta = \pm 1$.*

From the following steps we discover that the MAP estimator is a minimum distance decision rule, i.e., it selects the value $\hat{\theta}$ as that value of θ which is closest to the measured value X:

$$\hat{\theta}_{\text{MAP}} = \text{argmax}_{\theta=1,-1} f(X|\theta)$$

$$= \text{argmin}_{\theta=1,-1}\{(X-\theta)^2\}$$

$$= \begin{cases} 1, & X \geq 0 \\ -1, & X < 0 \end{cases}$$

On the other hand, the CME estimator is

$$\begin{aligned} \hat{\theta}_{CME} &= (1)P(\theta = 1|X) + (-1)P(\theta = -1|X) \\ &= \frac{\exp\left(-\frac{(X-1)^2}{2\sigma^2}\right) - \exp\left(-\frac{(X+1)^2}{2\sigma^2}\right)}{\exp\left(-\frac{(X-1)^2}{2\sigma^2}\right) + \exp\left(-\frac{(X+1)^2}{2\sigma^2}\right)}. \end{aligned}$$

The MAP and CME estimators are illustrated in Fog. 15. Unfortunately, we cannot derive the CmE since it is not well defined for discrete valued parameters θ (why?).

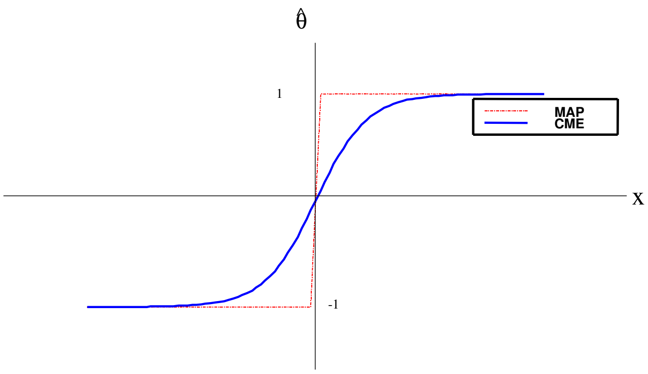


Figure 15: *MAP (light-font sign function) estimator and CME (heavy-font "S" curve) as functions of the measurement x. Only the MAP estimator gives the correct discrete range of values $\{-1, 1\}$ for θ*

Based on these above examples we make the summary remarks:

1. Different error criteria usually give different optimal estimators.

2. Optimal estimators of random parameters are not invariant to functional transformations. Specifically, if $\widehat{g(\theta)}$ is an optimal estimator of $g(\theta)$ and $\hat{\theta}$ is an optimal estimator of θ:

$$\widehat{g(\theta)} \neq g(\hat{\theta})$$

in general.

3. When they exist, the CmE and MAP estimators always take values in the parameter space Θ. The values taken by CME may fall outside of Θ, e.g., if it is discrete or if it is not a convex set.

4. The "MAP equation" stationary point condition $\partial f(\theta|x)/\partial\theta = 0$ at $\theta = \hat{\theta}_{MAP}$ is only useful for continuous densities that are differentiable and concave in continuous valued parameters θ (Fig. 16).

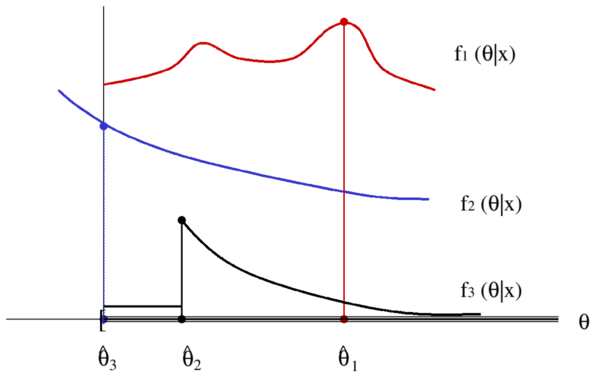


Figure 16: *Use of the stationary point MAP equation can fail to find the MAP estimator. In general there may exist no stationary points of the posterior density (f_2, f_3). or there may be multiple stationary points of the posterior density (f_1).*

4.3 ESTIMATION OF RANDOM VECTOR VALUED PARAMETERS

Define a vector parameter $\underline{\theta} \in \Theta \subset \mathbb{R}^p$, $\underline{\theta} = [\theta_1, \ldots, \theta_p]^T$, and define the s-norm on Θ

$$\|\underline{\theta}\|_s = \left(\sum_{i=1}^p |\theta_i|^s\right)^{1/s}.$$

Note that when $s = \infty$ this norm is equal to the maximum of the $|\theta_i|$'s.

The previously introduced scalar estimation criterion $E[c(\hat{\theta}, \theta)]$ needs to be generalized to handle vector parameter estimation. This turns out to be quite easy, at least for two of our proposed estimation criteria. Some possible generalizations of the previous three scalar criteria are (Figs. 17-20)

Estimator MSE:

$$\mathrm{MSE}(\hat{\underline{\theta}}) = E[\|\hat{\underline{\theta}} - \underline{\theta}\|_2^2] = \sum_{i=1}^p E[(\hat{\theta}_i - \theta_i)^2].$$

Estimator MAE:

$$\mathrm{MAE}(\hat{\theta}) = E[\|\hat{\theta} - \theta\|_1] = \sum_{i=1}^p E[|\hat{\theta}_i - \theta_i|].$$

Estimator Error Probability (MUE) - $(0 < p < \infty)$:

$$P_e(\hat{\theta}) = 1 - P(\|\hat{\theta} - \theta\|_p \leq \epsilon).$$

When $p = \infty$ P_e is the probability that the magnitude of at least one element of the vector $\hat{\theta} - \theta$ exceeds ϵ.

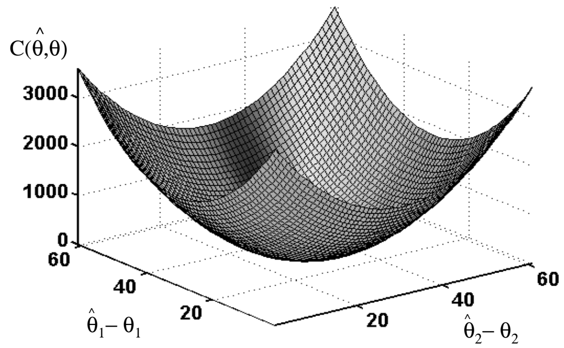


Figure 17: *Squared error criterion*

The MAE criterion, also known as total variation norm, does not often lead to unique optimal vector-valued estimators. Although the total variation norm has been of substantial recent interest, in our introductory treatment only MSE and P_e will be discussed.

4.3.1 VECTOR SQUARED ERROR

As $\text{MSE}(\hat{\theta}) = \sum_{i=1}^{p} \text{MSE}(\hat{\theta}_i)$ is an additive function, the minimum MSE vector estimator attains the minimum of each component $\text{MSE}(\hat{\theta}_i)$, $i = 1, \ldots, p$. Hence, we have the nice result that the vector minimum MSE estimator is simply the vector of scalar CME's for each component:

$$\underline{\hat{\theta}}_{\text{CME}} = E[\underline{\theta}|X] = \begin{bmatrix} E[\theta_1|X] \\ \vdots \\ E[\theta_p|X] \end{bmatrix}$$

As in the case of scalar estimation the minimum MSE estimator is the center of mass of the multivariate posterior density (Figs. 21-22).

4.3.2 VECTOR UNIFORM ERROR

For small ϵ the minimum mean uniform error (P_e) is attained by the vector MAP estimator which has form similar to the scalar MAP estimator

$$\underline{\hat{\theta}}_{\text{MAP}} = \text{argmax}_{\underline{\theta} \in \Theta} f(\underline{\theta}|x).$$

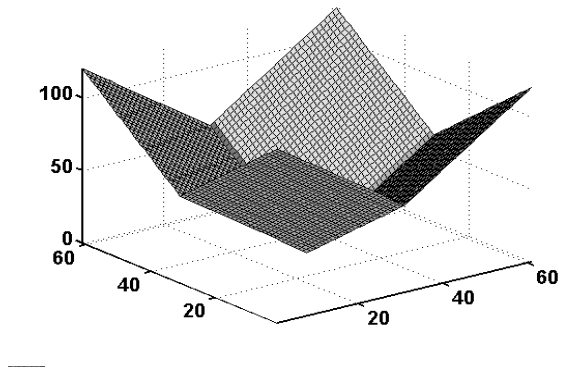


Figure 18: *Absolute error criterion*

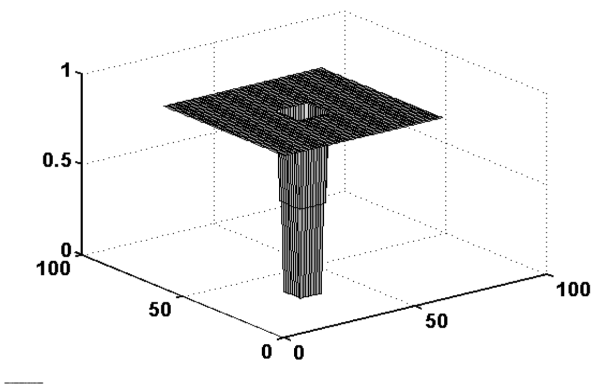


Figure 19: *Uniform error criterion*

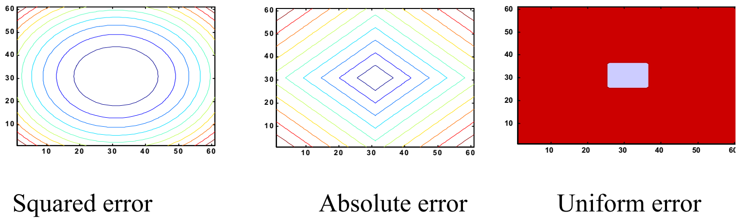


Figure 20: *Constant contours of three error criteria*

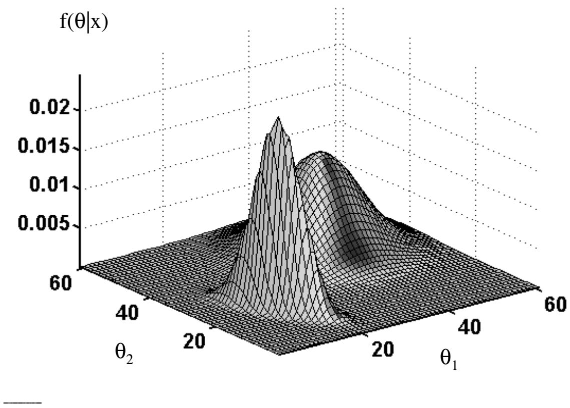


Figure 21: *Bivariate posterior density of two unknown parameters. Optimal estimates shown in Fig. 22.*

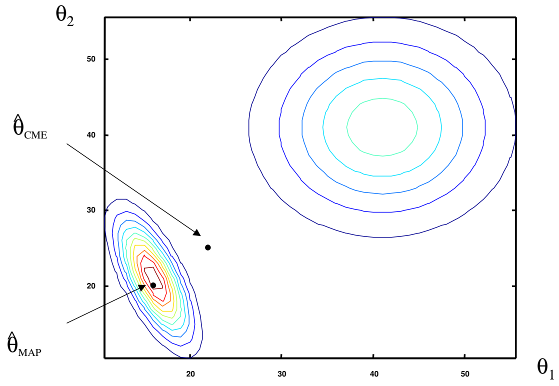


Figure 22: *Vector MAP estimate and CME for bivariate posterior illustrated in Fig. 23. The MAP estimate occurs at the global maximum of the posterior while the CME occurs at the center of mass.*

4.4 ESTIMATION OF NON-RANDOM PARAMETERS

To estimate random parameters one has a prior distribution and we can define a global estimation error criterion, the Bayes risk, which depends on the prior but not on any particular value of the parameter. In non-random parameter estimation there is no prior distribution. One can of course look at the problem of estimation of non-random parameters as estimation of random parameters conditioned on the value of the parameter, which we could call the true value. However, the formulation of optimal non-random parameter estimation requires a completely different approach. This is because if we do not have a prior distribution on the parameter virtually any reasonable estimation error criterion will be local, i.e., it will depend on the true parameter value. Thus we will need to define weaker properties than minimum risk, such as unbiasedness, that a good estimator of non-random parameters should have.

As before we first consider estimation of scalar non-random parameters θ. In this case it does not make sense to use the conditional density notation $f(\underline{x}|\theta)$ and we revert to the alternative notation for the model $f_\theta(\underline{x}) = f(\underline{x};\theta)$.

So, what are some possible design criteria for estimators of scalar real θ? One could try to minimize MSE, defined as

$$\mathrm{MSE}_\theta = E_\theta[(\hat{\theta} - \theta)^2].$$

Here we encounter a difficulty: if the true value θ is θ_0, the constant estimator $\hat{\theta} = c$ attains 0 MSE when $\theta_o = c$ (Fig. 23).

4.4.1 SCALAR ESTIMATION CRITERIA FOR NON-RANDOM PARAMETERS

Some possible scalar criteria for designing good estimators are the minimax criteria.

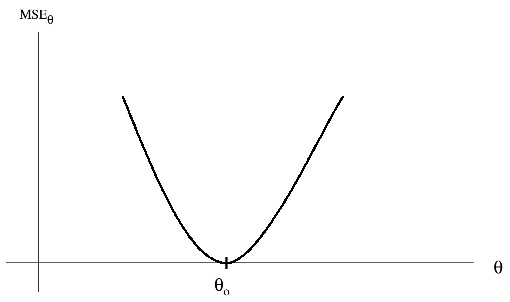


Figure 23: *MSE curve as function of θ for trivial estimator $\hat{\theta} = \theta_o$ of non-random parameter.*

1. Minimize worst case MSE. Choose $\hat{\theta}$ to minimize

$$\max_\theta \mathrm{MSE}_\theta(\hat{\theta}) = \max_\theta E_\theta[(\hat{\theta} - \theta)^2]$$

2. Minimize worst case estimator error probability:

$$\max_\theta P_e = \max_\theta P_\theta(|\hat{\theta} - \theta| > \epsilon)$$

If we would be satisfied by minimizing an upper bound on $\max P_e$, then we could invoke Tchebychev inequality

$$P_\theta(|\hat{\theta} - \theta| \geq \epsilon) \leq \frac{E_\theta[|\hat{\theta} - \theta|^2]}{\epsilon^2} \tag{35}$$

and focus on minimizing the worst case MSE. There is a large literature on minimax MSE estimation, see for example [40], but the mathematical level necessary to develop this theory is too advanced for an introductory treatment. We will not consider minimax estimation further in this book.

We next give several weaker conditions that a good estimator should satisfy, namely consistency and unbiasedness.

Definition: $\hat{\theta}_n = \hat{\theta}(X_1, \ldots, X_n)$ is said to be (weakly) *consistent* if for all θ and all $\epsilon > 0$

$$\lim_{n\to\infty} P_\theta(|\hat{\theta}_n - \theta| > \epsilon) = 0$$

This means that $\hat{\theta}_n$ converges in probability to the true parameter θ. It also means that the pdf of the estimator concentrates about θ (Fig. 24). Furthermore, by the Tchebychev inequality (35), if MSE goes to zero as $n \to \infty$ then $\hat{\theta}_n$ is consistent. As the MSE is usually easier to derive than

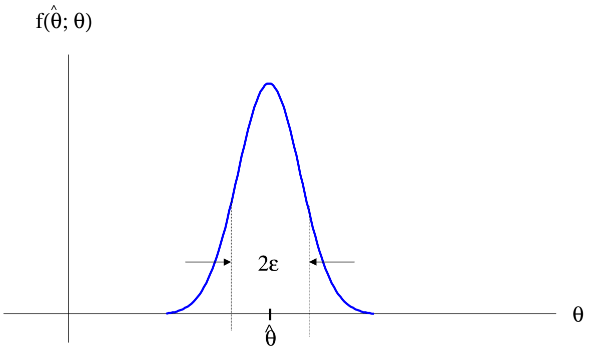


Figure 24: *Density $f(\hat{\theta};\theta)$ of $\hat{\theta}$ measures concentration of $\hat{\theta}$ about true parameter θ*

P_e, showing that MSE converges to zero is the typical way that one shows that an estimator is consistent.

For an estimator $\hat{\theta}$ define the *estimator bias* at a point θ to be

$$b_\theta(\hat{\theta}) = E_\theta[\hat{\theta}] - \theta.$$

Likewise the estimator variance is

$$\mathrm{var}_\theta(\hat{\theta}) = E_\theta[(\hat{\theta} - E_\theta[\hat{\theta}])^2].$$

Here the reader should recall the definition of the expectation operator E_θ: $E_\theta[g(X)] = \int_{\mathcal{X}} g(x)f(x;\theta)dx$, where X is a r.v. with density $f(x;\theta)$. As compared to the Bayes expectation $E[g(X)]$ used for random parameters, this expectation acts like a conditional expectation given a specific value of θ.

It is natural to require that a good estimator be *unbiased*, i.e., $b_\theta(\hat{\theta}) = 0$ for all $\theta \in \Theta$. This suggests a reasonable design approach: constrain the class of admissible estimators to be unbiased and try to find one that minimizes variance over this class. In some cases such an approach leads to a really good, in fact optimal, unbiased estimator called a UMVU estimator (Fig. 25). A caveat to the reader is necessary however: there exist situations where unbiasedness is not a desirable property to impose on an estimator. For example there are models for which no unbiased estimator of the model parameter exists and others for which the biased estimator has unreasonably high MSE, see Exercises at the end of this chapter and [58, Sec. 7.11, 7.15]. Fortunately, such models do not frequently arise in signal processing applications.

Definition: $\hat{\theta}$ is said to be a uniform minimum variance unbiased (UMVU) estimator if for all $\theta \in \Theta$ it has less variance than any other unbiased estimator $\hat{\hat{\theta}}$. Thus a UMVU estimator satisfies

$$\mathrm{var}_\theta(\hat{\theta}) \le \mathrm{var}_\theta(\hat{\hat{\theta}}), \quad \theta \in \Theta$$

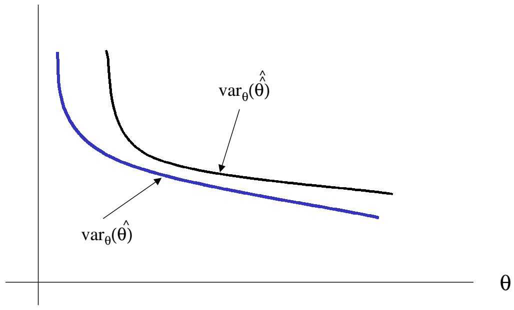


Figure 25: *A UMVU estimator $\hat{\theta}$ is an unbiased estimator that has lower variance than any other unbiased estimator $\hat{\hat{\theta}}$*

Unfortunately, UMVU estimators only rarely exist for finite number n of samples $X_1, \ldots, X_n$. Thus one is usually forced to sacrifice the unbiasedness constraint in order to develop good tractable estimation procedures. For such estimators there exists an important relation between MSE, variance and bias:

$$
\begin{aligned}
\mathrm{MSE}_\theta(\hat{\theta}) &= E_\theta[(\hat{\theta}-\theta)^2] = E_\theta[\Big((\hat{\theta}-E_\theta[\hat{\theta}]) + (E_\theta[\hat{\theta}]-\theta)\Big)^2] \\
&= \underbrace{E_\theta[(\hat{\theta}-E_\theta[\hat{\theta}])^2]}_{\mathrm{var}_\theta(\hat{\theta})} + \underbrace{\Big(E_\theta[\hat{\theta}]-\theta\Big)^2}_{b_\theta(\hat{\theta})} + 2\underbrace{E_\theta[\hat{\theta}-E_\theta[\hat{\theta}]]}_{=0}\ b_\theta(\hat{\theta}) \\
&= \mathrm{var}_\theta(\hat{\theta}) + b_\theta^2(\hat{\theta})
\end{aligned}
$$

The above relation implies that in general, for specified MSE, there always exists a "bias-variance tradeoff," at least for good estimators: any reduction in bias comes at the expense of an increase in variance.

We now get down to the business of defining some general procedures for designing good estimators of non-random parameters. Two important classes of estimation procedures we will consider are:

* method of moments

* maximum likelihood

4.4.2 METHOD OF MOMENTS (MOM) SCALAR ESTIMATORS

The method of moments is a very natural procedure which consists in finding the parameter that attains the best match between empirically computed moments and ensemble moments. Specifically, for positive integer k let $m_k = m_k(\theta)$ be the k-th order ensemble moment of $f(x;\theta)$:

$$m_k = E_\theta[X^k] = \int x^k f(x;\theta)dx.$$

What if we could find a set of K moments such that some vector function $\underline{h}$ could be found that satisfies

$$\theta = \underline{h}(m_1(\theta), \ldots, m_K(\theta)).$$

For example, let's say we could compute a closed form expression $g(\theta)$ for the k-th ensemble moment $E_\theta[X^k]$ and found that the function g was invertible. Then if someone only reported the value m_k of this ensemble moment without specifying the θ for which it was computed we could recover θ by applying the inverse function

$$\theta = g^{-1}(m_k).$$

Since g^{-1} recovers θ from the ensemble moment of X, if we only have access to an i.i.d. sample $X_1, \ldots, X_n$ from $f(x;\theta)$ it makes sense to estimate θ by applying g^{-1} to an estimated moment such as the empirical average

$$\hat{m}_k = \frac{1}{n}\sum_{i=1}^n X_i^k,$$

yielding the estimator

$$\hat{\theta} = g^{-1}(\hat{m}_k).$$

In many cases it is difficult to find a single ensemble moment that gives an invertible function of θ. Indeed, using only the k-th moment we may only be able to find a constraint equation $g(\theta) = \hat{m}_k$ that gives several possible solutions $\hat{\theta}$. In these cases, one can sometimes compute other ensemble and empirical moments to construct more constraint equations and force a unique solution. We will explore this approach in the examples below. Next we give some important asymptotic optimality properties of MOM estimators (see Serfling [62] for proofs).

IMPORTANT PROPERTIES OF MOM ESTIMATORS

When the moments m_k are smooth functions of the parameter θ and an inverse function g^{-1}, described above, exists:

1. MOM estimators are asymptotically unbiased as $n \to \infty$
2. MOM estimators are consistent

Note that MOM estimators are not always unbiased in the finite sample regime. There are, however, some inherent difficulties that one sometimes encounters with MOM which are summarized below.

1. MOM estimator is not unique, i.e., it depends on what order moment is used.
2. MOM is inapplicable in cases where moments do not exist (e.g. Cauchy p.d.f.) or are unstable.

An alternative to MOM which can sometimes circumvent the existence problem is to match sample and ensemble fractional moments m_k where k is a positive rational number less than one. Fractional moments can exist when integer moments do not exist and can be quite useful in these situations [63].

Let's do some examples.

Example 13 $\underline{X}$ *i.i.d. Bernoulli random variables*

Bernoulli measurements arise anytime one deals with (binary) quantized versions of continuous variables, e.g., thresholded radar signals ("radar return is above or below a threshold"), failure data, or digital media, e.g., Internet measurements. In these cases the parameter of interest is typically the probability of success, i.e., the probability that the measured variable is a "logical 1."

The model is that $\underline{X} = [X_1, \ldots, X_n]$ are i.i.d. with

$$X_i \sim f(x;\theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1.$$

Here $\theta \in [0,1]$ or, more specifically, $\theta = P(X_i = 1)$, $1-\theta = P(X_i = 0)$.

Objective: find a MOM estimator of θ

Note that for any $k > 0$ $E[X_i^k] = P(X_i = 1) = \theta$ so that all moments are identical and the function g mapping moments to θ is the identity map. Thus a MOM estimator of θ is simply sample mean:

$$\hat{\theta} = \overline{X}.$$

It is obvious that $\hat{\theta}$ is unbiased since $E_\theta[\overline{X}] = m_1 = \theta$. Furthermore, it has variance taking a maximum at $\theta = \frac{1}{2}$ (Fig. 26)

$$\mathrm{var}_\theta(\overline{X}) = (m_2 - m_1^2)/n = \theta(1-\theta)/n.$$

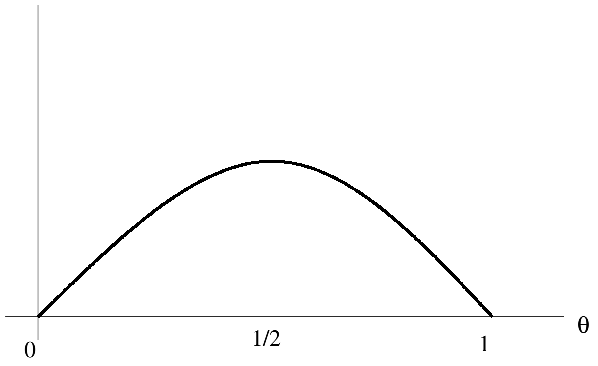


Figure 26: *Variance of MOM estimator of probability of success of Bernoulli r.v.*

Reiterating, for this Bernoulli example the order of the moment used in the moment matching process leads to identical MOM estimators. This behavior of MOM is very unusual.

Example 14 $\underline{X}$ *i.i.d. Poisson random variables*

Poisson measurements are ubiquitous in many scenarios where there are counting measurements. For example, in positron emission tomography (PET) the decay of an isotope in a particular spatial location within a patient's body produces a gamma ray which is registered as a single "count" on a detector. The temporal record of the times at which these counts are registered on the detector forms a Poisson process [65]. The total number of counts registered over a finite time interval is a Poisson random variable with rate parameter determined by the mean concentration of the isotope. The objective of a PET system is to reconstruct, i.e., estimate, the distribution of rates over the imaging volume. The Poisson distribution is also frequently used as a model for the number of components or degrees of freedom generating the measured values. For example, the number of molecules in a mass spectroscopy measurement, the number of atoms in a molecule, or the number of targets in a cell detected by a radar.

Again assuming i.i.d. measurements, the model for each data sample is

$$X_i \sim p(x;\theta) = \frac{\theta^x}{x!}e^{-\theta}, \quad x = 0, 1, 2, \ldots,$$

where $\theta > 0$ is the unknown rate. It is readily verified that the mean m_1 is equal to θ. Therefore, like in the Bernoulli example a MOM estimator of θ is the sample mean

$$\hat{\theta}_1 = \overline{X}.$$

Alternatively, as the second moment satisfies $m_2 = \theta + \theta^2$, another MOM estimator is the (positive) value of $\hat{\theta}_2$ which satisfies the equation : $\hat{\theta}_2 + \hat{\theta}_2^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 := \overline{X^2}$, i.e.

$$\hat{\theta}_2 = \frac{-1 \pm \sqrt{1 + 4\overline{X^2}}}{2}.$$

As yet another example, we can express m_2 as $m_2 = \theta + m_1^2$ or $\theta = m_2 - m_1^2 = \mathrm{var}_\theta(X_i)$. Hence, a MOM estimator is

$$\hat{\theta}_3 = \overline{X^2} - \overline{X}^2 = n^{-1}\sum_{i=1}^n (X_i - \overline{X})^2.$$

Among all of these MOM estimators only the sample mean estimator is unbiased for finite n:

$$\begin{aligned} E_\theta(\hat{\theta}_1) &= \theta, & \mathrm{var}_\theta(\hat{\theta}_1) &= \theta/n, \\ E_\theta(\hat{\theta}_3) &= \frac{n-1}{n}\theta, & \mathrm{var}_\theta(\hat{\theta}_3) &\approx (2\theta^2 + \theta)/n. \end{aligned}$$

Closed form expressions for bias and variance of $\hat{\theta}_2$ do not exist.

You should notice that $\hat{\theta}_1$ compares favorably to $\hat{\theta}_3$ since it has both lower bias and lower variance.

We make the following observations.

1. $\hat{\theta}_1$ is unbiased for all n.
2. $\hat{\theta}_2, \hat{\theta}_3$ are asymptotically unbiased as $n \to \infty$.
3. Consistency of $\hat{\theta}_1$ and $\hat{\theta}_3$ is directly verifiable from the above expressions for mean and variance and Thebychev's inequality.

4.4.3 MAXIMUM LIKELIHOOD (ML) SCALAR ESTIMATORS

Maximum likelihood (ML) is arguably the most commonly adopted parametric estimation principle in signal processing. This is undoubtedly due to the fact that, unlike other methods, ML usually results in unique estimators and is straightforward to apply to almost all problems.

For a measurement $\underline{X} = \underline{x}$ we define the "likelihood function" for θ

$$L(\theta) = f(\underline{x}; \theta)$$

and the log-likelihood function

$$l(\theta) = \ln f(\underline{x}; \theta).$$

These should be viewed as functions of θ for a fixed value of $\underline{x}$ (Fig. 27). Readers may find it strange that the $\underline{x}$-dependence of the functions $L(\theta)$ and $l(\theta)$ is not indicated explicitly. This convention of dropping such dependencies to clarify the "working" variable θ is common in statistics and signal processing.

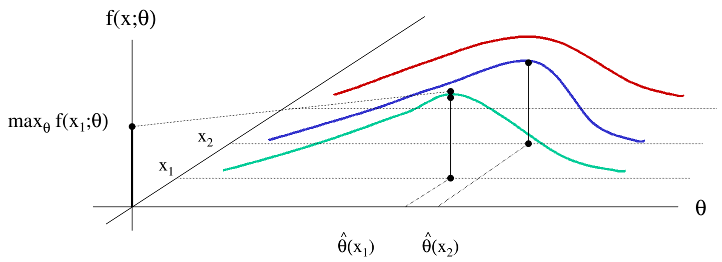


Figure 27: *The likelihood function for θ*

The ML estimator $\hat{\theta}$ is defined as the value of θ which causes the data $\underline{x}$ to become "most likely," i.e., $\hat{\theta}$ makes it most likely that $\underline{x}$ was generated from $f(\underline{x}; \theta)$. Mathematically, we have the equivalent definitions

$$\begin{aligned}
\hat{\theta} &= \text{argmax}_{\theta \in \Theta} f(\underline{X}; \theta) \\
&= \text{argmax}_{\theta \in \Theta} L(\theta) \\
&= \text{argmax}_{\theta \in \Theta} l(\theta).
\end{aligned}$$

In fact the ML estimate can be found by maximizing any monotone increasing function of $L(\theta)$.

Important properties of ML estimators for smooth likelihoods (Ibragimov and Has'minskii [29], Serfling [62]) are

Property 1. MLE's are asymptotically unbiased. The proof requires additional technical conditions.

Property 2. MLE's are consistent. The proof requires additional technical conditions.

Property 3. Unlike many other estimators, e.g. MAP and UMVUE estimators, MLE's are invariant to any transformation of the parameters, i.e.,

$$\varphi = g(\theta) \quad \Rightarrow \quad \hat{\varphi} = g(\hat{\theta}).$$

This is easy to see for monotone transformations (Fig. 28) but in fact it applies to arbitrary transformations (See exercises).

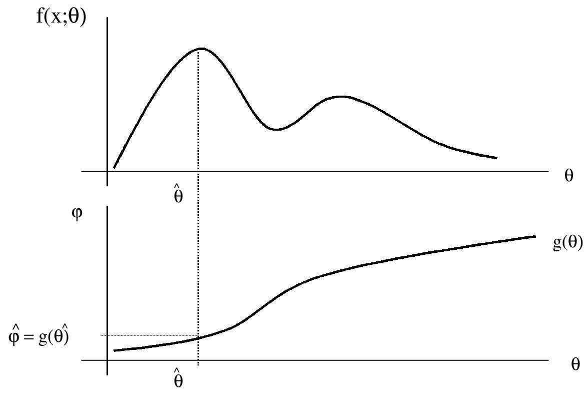


Figure 28: *Invariance of MLE to functional transformation g*

Property 4: MLE's are asymptotically UMVU in the sense that

$$\lim_{n\to\infty} n\mathrm{var}_\theta(\hat{\theta}) = \frac{1}{F_1(\theta)},$$

where F_1 is a quantity known as the Fisher information, which will be introduced soon, and $1/F_1$ specifies the fastest possible asymptotic rate of decay of any unbiased estimator's variance. The proof requires additional technical conditions.

Property 5: MLE's are asymptotically Gaussian in the sense

$$\sqrt{n}(\hat{\theta}_n - \theta) \to Z, \quad (i.d.)$$

where $Z \sim \mathcal{N}(0, 1/F_1(\theta))$. Here the notation i.d. denotes convergence *in distribution*. This means that the cumulative distribution function (cdf) of $\sqrt{n}(\hat{\theta}_n - \theta)$ converges to the (standard normal) cdf of Z. The proof requires additional technical conditions.

Property 6: The MLE is equivalent to the MAP estimator for a uniform prior $f(\theta) = c$.

Property 7: If the MLE is unique, the MLE is a function of the data only through the sufficient statistic.

Now let's go back and revisit our MOM examples with the MLE in mind.

Example 15 $\underline{X}$ *i.i.d. Bernoulli random variables*

We can solve for the MLE in two ways: (1) considering the entire observation $\underline{X}$; and (2) considering only a sufficient statistic $T(\underline{X})$.

1. With the entire observation $\underline{X} = \underline{x}$ the likelihood function is the product

$$L(\theta) = f(\underline{x}; \theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}.$$

   Is is convenient to rewrite this in the form

$$\begin{aligned} L(\theta) &= \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} \\ &= \theta^{n\overline{x_i}}(1-\theta)^{n-n\overline{x_i}}. \end{aligned} \qquad (36)$$

   As this function smooth and concave in θ, differentiation with respect to θ yields a stationary point condition, the "ML equation," for the MLE $\hat{\theta}$

$$0 = \frac{\partial}{\partial \hat{\theta}} f(\underline{x}; \hat{\theta}) = n \left[\frac{(1-\hat{\theta})\overline{x_i} - \hat{\theta}(1-\overline{x_i})}{\hat{\theta}(1-\hat{\theta})} \right] f(\underline{x}; \hat{\theta}).$$

   Solving the equation $(1-\hat{\theta})\overline{x_i} - \hat{\theta}(1-\overline{x_i}) = 0$ we obtain the MLE

$$\hat{\theta} = \overline{X}, \qquad (37)$$

   which is identical to the MOM estimator obtained above.

2. Using the Fisher factorization (24) on the p.d.f. (36) of $\underline{X}$ it is easily seen that $T(\underline{X}) = \sum_{i=1}^n X_i$ is a sufficient statistic for θ. The distribution of T is *binomial* with parameter θ:

$$f_T(t; \theta) = \binom{n}{t} \theta^t (1-\theta)^{n-t}, \quad t = 0, \ldots, n,$$

   where the subscript T on the density of T is to clarify that this is the p.d.f. of the r.v. T. Identification of $t = n\overline{X}$ reveals that this is of exactly the same form, except for a constant multiplication factor, as (36). The ML equation is therefore the same as before and we obtain the identical MLE estimator (37).

Example 16 $\underline{X}$ *i.i.d. Poisson random variables*

To find the MLE of the rate parameter θ express the density of the samples as:

$$f(\underline{x}; \theta) = \prod_{i=1}^{n} \frac{\theta^{x_i}}{x_i!} e^{-\theta}.$$

The likelihood function $L(\theta) = f(\underline{x};\theta)$ has to be maximized over θ to produce the MLE. It is more convenient to deal with the log likelihood

$$\hat{\theta}_{ml} = \text{argmax}_{\theta>0} \ln L(\theta)$$

and we have

$$\begin{aligned}
l(\theta) &= \ln f(\underline{x};\theta) \\
&= \ln \prod_{k=1}^{n} \frac{\theta^{x_k}}{x_k!} e^{-\theta} \\
&= \sum_{k=1}^{n} x_k \ln\theta - n\theta - \underbrace{\sum_{k=1}^{n} \ln x_k!}_{\text{constant in } \theta} \\
&= \overline{x_i} n \ \ln\theta - n\theta + c,
\end{aligned}$$

where c is an irrelevant constant.

It is easily verified (look at second derivative) that the log-likelihood $l(\theta)$ is a smooth strictly concave function of θ. Thus the MLE is the unique solution $\theta = \hat{\theta}$ of the equation

$$0 = \partial \ln f/\partial\theta = \frac{n\overline{x_i}}{\theta} - n.$$

We find that the MLE is identical to the first MOM estimator we found for this problem:

$$\hat{\theta} = \overline{X},$$

which we know is unbiased with variance equal to θ.

Let's check the asymptotic Gaussian property. Write

$$\begin{aligned}
\sqrt{n}(\overline{X} - \theta) &= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \theta)\right) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n}(X_i - \theta).
\end{aligned}$$

By the central limit theorem (CLT), this converges in distribution to a Gaussian r.v.

$$\begin{aligned}
E_\theta[\sqrt{n}(\overline{X} - \theta)] &= 0 \\
\text{var}_\theta(\sqrt{n}(\overline{X} - \theta)) &= \theta.
\end{aligned}$$

4.4.4 SCALAR CRAMÈR-RAO BOUND (CRB) ON ESTIMATOR VARIANCE

The CRB can be defined for both random and non-random parameters. However the CRB is more useful for non-random parameters as it can be used to establish optimality or near optimality of an unbiased candidate estimator. Unlike the non-random case, for random parameters the optimal estimator and its MSE are functions of the known joint density of θ and $\underline{X}$. Thus there exist more accurate alternatives to the CRB for approximating estimator MSE, most of which boil down to approximating an integral representation of the minimum mean squared error. We therefore focus our energies on the CRB for non-random parameters - the interested reader can refer to [73] for the random case.

The Cramèr-Rao Lower Bound *Let $\theta \in \Theta$ be a non-random scalar and assume:*

1. Θ is an open subset, e.g. (a,b), of $\mathbb{R}$.

2. $f(\underline{x};\theta)$ is smooth (Ibragimov and Has'minskii [29]) and differentiable in θ.

The following is the Cramèr-Rao bound for scalar θ

For any unbiased estimator $\hat{\theta}$ of θ

$$\mathrm{var}_\theta(\hat{\theta}) \geq 1/F(\theta), \quad , \tag{38}$$

where "=" is attained iff for some non-random scalar k_θ

$$\frac{\partial}{\partial\theta} \ln f(\underline{x};\theta) = k_\theta(\hat{\theta} - \theta). \tag{39}$$

Here k_θ is a constant that can depend on θ but not on x. When the CRB is attainable it is said to be a tight bound and (39) is called the CRB tightness condition.

In the CRB $F(\theta)$ is the Fisher information which can be shown [73] to take on either of the following two equivalent forms:

$$\begin{aligned} F(\theta) &= E_\theta\left[\left(\frac{\partial}{\partial\theta} \ln f(\underline{X};\theta)\right)^2\right] \\ &= -E_\theta\left[\frac{\partial^2}{\partial\theta^2} \ln f(\underline{X};\theta)\right] \end{aligned}$$

This latter second derivative form of the Fisher information can be used to show that the scalar k_θ in the tightness condition (39) is in fact equal to $F(\theta)$. To see this simply differentiate both sides of the equation (39), take expectations, and use the fact that $\hat{\theta}$ is unbiased.

Before going on to some examples, we provide a simple derivation of the scalar CRB here. A more detailed proof of the more general vector parameter CRB will be given later. There are three steps to the derivation of the scalar CRB - assuming that interchange of the order of integration and differentiation is valid. The first step is to notice that the mean of the derivative of the log-likelihood is equal to zero:

$$\begin{aligned} E_\theta[\partial \ln f_\theta(\underline{X})/\partial\theta] &= E_\theta\left[\frac{\partial f_\theta(\underline{X})/\partial\theta}{f_\theta(\underline{X})}\right] \\ &= \int \frac{\partial}{\partial\theta} f_\theta(\underline{x}) d\underline{x} \\ &= \frac{\partial}{\partial\theta} \underbrace{\int f_\theta(\underline{x})\, d\underline{x}}_{=1} \\ &= 0 \end{aligned}$$

The second step is to show that the correlation between the derivative of the log-likelihood and the estimator is a constant:

$$\begin{aligned} E_\theta[(\hat{\theta}(\underline{X}) - E_\theta[\hat{\theta}])(\partial \log f_\theta(\underline{X})/\partial\theta)] &= \int (\hat{\theta}(\underline{x}) - E_\theta[\hat{\theta}]) \frac{\partial}{\partial\theta} f_\theta(\underline{x}) d\underline{x} \\ &= \frac{\partial}{\partial\theta} \underbrace{\int \hat{\theta}(\underline{x}) f_\theta(\underline{x})\, d\underline{x}}_{=E_\theta[\hat{\theta}]=\theta} \\ &= 1 \end{aligned}$$

Where we have used the result of step 1 in line 2 above. Finally, apply the Cauchy-Schwarz (CS) inequality $E^2[UV] \leq E[U^2]E[V^2]$ to obtain:

$$\begin{aligned} 1 &= E_\theta^2[(\hat{\theta}(\underline{X}) - E_\theta[\hat{\theta}])(\partial \ln f_\theta(\underline{X})/\partial\theta)] \\ &\leq E_\theta[(\hat{\theta}(\underline{X}) - E_\theta[\hat{\theta}])^2] \cdot E_\theta[(\partial \ln f_\theta(\underline{X})/\partial\theta)^2] \\ &= \mathrm{var}_\theta(\hat{\theta}) \cdot F(\theta). \end{aligned}$$

Equality occurs in the CS inequality if and only if $U = kV$ for some non-random constant k. This gives (38) and completes the derivation of the CRB.

To illustrate the CRB let's go back and reconsider one of the previous examples.

Example 17 *CRB for the Poisson rate*

Assume again that $\underline{X} = [X_1, \ldots, X_n]$ is a vector of i.i.d. Poisson random variables

$$X_i \sim f(x;\theta) = \frac{\theta^x}{x!} e^{-\theta}, \quad x = 0, 1, 2, \ldots$$

To find the CRB we must first compute the Fisher information. Start with

$$\ln f(\underline{x};\theta) = \sum_{k=1}^n x_k \ln\theta - n\theta - \underbrace{\sum_{k=1}^n \ln x_k!}_{\text{constant in } \theta},$$

and differentiate twice

$$\partial \ln f(\underline{x};\theta)/\partial\theta = \frac{1}{\theta}\sum_{k=1}^n x_k - n \tag{40}$$

$$\partial^2 \ln f(\underline{x};\theta)/\partial\theta^2 = -\frac{1}{\theta^2}\sum_{k=1}^n x_k. \tag{41}$$

Therefore, as $E[\sum_{k=1}^n X_k] = n\theta$, the Fisher information given the n i.i.d. samples is

$$F_n(\theta) = \frac{n}{\theta}.$$

The CRB asserts that for any unbiased estimator of the Poisson rate θ

$$\mathrm{var}_\theta(\hat{\theta}) \geq \frac{\theta}{n}.$$

It is useful to make the following key observations.

Observation 1: From example (14) we know that the sample mean $\overline{X}$ is unbiased and has $\text{var}_\theta(\overline{X}) = \theta/n$. This is equal to the CRB and we conclude the CRB is tight.

Observation 2: In fact we could have concluded by inspection that the unbiased estimator $\overline{X}$ achieves the CRB; i.e., without having to explicitly compute its variance and compare to one over the Fisher information. This follows from the fact that equation (40) implies that the CRB tightness condition (39) is satisfied:

$$\partial \ln f(\underline{X};\theta)/\partial\theta = \frac{1}{\theta}\sum_{k=1}^n X_k - n = \underbrace{\frac{n}{\theta}}_{k_\theta}(\underbrace{\overline{X}}_{\hat{\theta}} - \theta). \tag{42}$$

Furthermore, once tightness is established in this fashion the variance of $\overline{X}$ can be computed by computing the CRB. This indirect method can sometimes be simpler than direct computation of estimator variance.

Observation 3: the expectation of the right hand side of (42) is zero since $\hat{\theta}$ is unbiased. This implies that

$$E_\theta\left[\partial \ln f(\underline{X};\theta)/\partial\theta\right] = 0.$$

The interpretation is that the gradient at θ of the log-likelihood is an unbiased estimator of zero when θ is the true parameter, i.e. the parameter appearing in the subscript of the expectation. This relation is generally true: it holds for any density satisfying the differentiability and smoothness conditions [29]) sufficient for existence of the CRB.

GENERAL PROPERTIES OF THE SCALAR CRB

Property 1. The Fisher information is a measure of the average (negative) curvature of the log likelihood function $\ln f(\underline{x};\theta)$ near the true θ (Kass and Voss [34]) (Fig. 30).

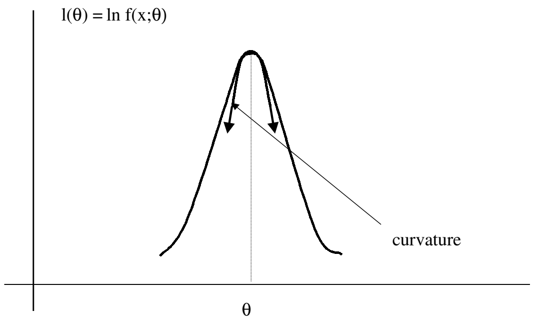


Figure 29: *The curvature of the log likelihood function* $\ln f(\underline{x};\theta)$ *in the vicinity of true* θ

Property 2. Let $F_n(\theta)$ be the Fisher information for a sample of n i.i.d. measurements $X_1, \ldots, X_n$. Then

$$F_n(\theta) = nF_1(\theta).$$

Hence, for smooth likelihood functions of continuous parameters, and unbiased estimators, the variance $\mathrm{var}_\theta(\hat{\theta})$ cannot decay faster than order $1/n$

Proof of Property 2:

Since $\underline{X} = [X_1, \ldots, X_n]^T$ are i.i.d.

$$f(\underline{x};\theta) = \prod_{i=1}^n f(x_i;\theta)$$

so that

$$\begin{aligned} F_n(\theta) &= -E\left[\frac{\partial^2}{\partial\theta^2} \ln f(\underline{X};\theta)\right] \\ &= -E\left[\sum_{i=1}^n \frac{\partial^2}{\partial\theta^2} \ln f(X_i;\theta)\right] \\ &= \sum_{i=1}^n \underbrace{-E\left[\frac{\partial^2}{\partial\theta^2} \ln f(X_i;\theta)\right]}_{F_1(\theta)} \end{aligned}$$

$\diamond$

For unbiased estimators, the CRB specifies an unachievable region of variance as a function of n (Fig. 30). Good unbiased estimators $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$ of scalar continuous parameters have variance that behaves as $\mathrm{var}_\theta(\hat{\theta}) = O(1/n)$.

Property 3. If $\hat{\theta}$ is unbiased and $\mathrm{var}_\theta(\hat{\theta})$ attains the CRB for all θ, $\hat{\theta}$ is said to be an *efficient estimator*. Efficient estimators are always UMVU (but not conversely, e.g., see counterexample in [58, Ch 9]). Furthermore, if an estimator is asymptotically unbiased and its variance decays with optimal rate constant

$$\lim_{n\to\infty} b_\theta(\hat{\theta}) = 0, \qquad \lim_{n\to\infty} n\mathrm{var}_\theta(\hat{\theta}) = 1/F_1(\theta),$$

where F_1 is the Fisher information given a single sample X_i, then $\hat{\theta}$ is said to be *asymptotically efficient.*

Exponential families play a special role with regard to efficiency. In particular, if X is a sample from a density in the exponential family with scalar parameter θ having the mean value parameterization (recall discussion in Sec. 3.5.4) then (See exercise 4.32)

$$\begin{aligned} \theta &= E_\theta[t(X)] & (43) \\ F(\theta) &= 1/\mathrm{var}_\theta(t(X)), & (44) \end{aligned}$$

where $F(\theta)$ is the Fisher information given the sample X. Therefore, if one has an i.i.d. sample $\underline{X} = [X_1, \ldots, X_n]^T$ from such a density then $\hat{\theta} = n^{-1}\sum_{i=1}^n t(X_i)$ is an unbiased and efficient estimator of θ.

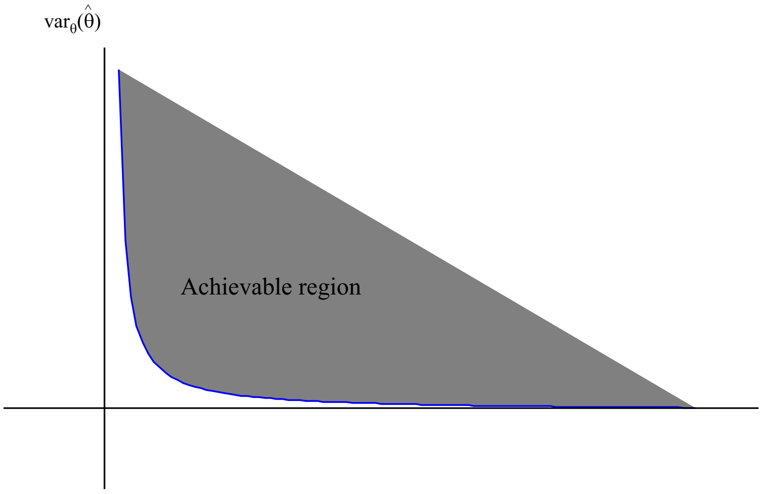


Figure 30: *The CRB defines an unachievable region of variance which is under the CRB curve, indicated by the unshaded area. Good unbiased estimators of continuous parameters have variance that decays as* $1/n$.

Somewhat surprisingly, the next property states that efficient estimators can exist only when the sample comes from an exponential family with mean value parameterization.

Property 4. Efficient estimators for θ can only exist when the underlying model is in an exponential family, defined in Sec. 3.5.4:

$$f(x;\theta) = a(\theta)b(x)e^{-c(\theta)t(x)}.$$

and when $E_\theta[t(X)] = \theta$, i.e., the density is in its mean value parameterization.

Proof of Property 4:

Without loss of generality we specialize to the case of a single sample $n = 1$ and $\Theta = (-\infty, \infty)$. Recall the condition for equality in the CR bound to be achieved by an estimator $\hat\theta$ is that the p.d.f. be expressible as

$$\frac{\partial}{\partial\theta} \ln f(x;\theta) = k_\theta(\hat\theta - \theta). \tag{45}$$

For fixed θ_o, integrate the LHS of condition (45) over $\theta \in [\theta_o, \theta']$

$$\int_{\theta_o}^{\theta'} \frac{\partial}{\partial\theta} \ln f(x;\theta)dx = \ln f(x;\theta') - \ln f(x;\theta_o).$$

On the other hand, integrating the RHS of the condition

$$\int_{\theta_o}^{\theta'} k_\theta(\hat\theta - \theta)d\theta \;\; = \;\; \hat\theta \underbrace{\int_{\theta_o}^{\theta'} k_\theta d\theta}_{c(\theta')} - \underbrace{\int_{\theta_o}^{\theta'} k_\theta \theta d\theta}_{d(\theta')}.$$

Or combining the integrals of RHS and LHS of (45)

$$f(x;\theta) = \underbrace{e^{-d(\theta)}}_{a(\theta)} \underbrace{f(x;\theta_o)}_{b(x)} e^{-c(\theta)} \overbrace{\hat{\theta}}^{t(x)} .$$

◇

We illustrate the above properties with two more examples.

Example 18 *Parameter estimation for the exponential density.*

A non-negative random variable X has an exponential density with mean θ if its p.d.f. is of the form $f(x;\theta) = \theta^{-1}\exp(-x/\theta)$ where $\theta > 0$. The exponential random variable is commonly used as a model for service time or waiting time in networks and other queuing systems. You can easily verify that this density is in the exponential family specified by $a(\theta) = \theta^{-1}$, $b(x) = I_{[0,\infty)}(x)$, $c(\theta) = -\theta^{-1}$ and $t(x) = x$. As $E_\theta[X] = \theta$ the p.d.f. $f(x;\theta)$ is in its mean value parametrization and we conclude that the sample mean $\overline{X}$ is an unbiased estimator of θ. Furthermore, it is efficient and therefore UMVU when n i.i.d. observations $\underline{X} = [X_1, \ldots, X_n]^T$ are available.

NOTE: we cannot conclude from the above arguments that $1/\overline{X}$ is an efficient estimator of $1/\theta$.

Example 19 $\underline{X}$ *i.i.d.,* $X_i \sim \mathcal{N}(\theta, \sigma^2)$

The Gaussian "bell curve" distribution arises in so many applications that it has become a standard model. Use of this model is usually justified by invocation of the Central Limit Theorem as describing the measurements, or measurement noise, as the sum of many small contributions, e.g. random atomic collisions, scattered light, aggregation of repeated measurements.

Our first objective will be to find the MLE and CRB for estimating the mean θ of univariate Gaussian with known variance σ^2. As the Gaussian with unknown mean is in the exponential family we could take the same approach as above to find efficient estimators. But let's spice things up and follow an alternative route of trying to tease an efficient estimator out of the tightness condition in the CRB.

$$f(\underline{x};\theta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^n (x_k - \theta)^2\right\}.$$

Or

$$\ln f(\underline{x};\underline{\theta}) = -\frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{k=1}^n (x_k - \theta)^2 + c,$$

where c is constant. Compute the first derivative

$$\begin{aligned} \partial \ln f/\partial\theta &= \frac{1}{\sigma^2}\sum_{k=1}^n (x_k - \theta) \\ &= \underbrace{\frac{n}{\sigma^2}}_{k_\theta} (\overline{x_i} - \theta). \end{aligned} \tag{46}$$

Thus the CRB tightness condition (39) is satisfied and we can identify, once again, the sample mean $\overline{x_i}$ as the optimal estimator of the common mean of a Gaussian sample.

We take another derivative of the log-likelihood with respect to θ and invert it to verify what we already knew about the variance of the sample mean

$$\mathrm{var}_\theta(\overline{X}) = 1/F_n(\theta) = \sigma^2/n.$$

The first inequality is only true since we know that $\overline{X}$ is efficient.

Note that the leading factor in the tight CRB condition (46) is: $k_\theta = \mathrm{var}_\theta^{-1}(\overline{X})$. This is always true for efficient estimators when k_θ does not depend on θ.

4.5 ESTIMATION OF MULTIPLE NON-RANDOM PARAMETERS

We now turn the more general problem of many unknown deterministic parameters. This problem is quite different from the previously studied case of multiple random parameters since there is no joint posterior density to marginalize. First we arrange all unknown parameters in a vector:

$$\underline{\theta} = [\theta_1, \ldots, \theta_p]^T,$$

and state the problem as finding a vector valued estimator $\hat{\underline{\theta}}$ of $\underline{\theta}$.

The joint density for the measurements $\underline{X}$ is written as:

$$f(\underline{x}; \theta_1, \ldots, \theta_p) = f(\underline{x}; \underline{\theta}).$$

POSSIBLE ESTIMATOR PERFORMANCE CRITERIA

As for a scalar estimator we define the vector estimator bias vector:

$$b_{\underline{\theta}}(\hat{\underline{\theta}}) = E_{\underline{\theta}}[\hat{\underline{\theta}}] - \underline{\theta},$$

and the symmetric estimator covariance matrix:

$$\begin{aligned}
\mathrm{cov}_{\underline{\theta}}(\hat{\underline{\theta}}) &= E_{\underline{\theta}}[(\hat{\underline{\theta}} - E[\hat{\underline{\theta}}])(\hat{\underline{\theta}} - E[\hat{\underline{\theta}}])^T] \\
&= \begin{bmatrix}
\mathrm{var}_{\underline{\theta}}(\hat{\theta}_1) & \mathrm{cov}_{\underline{\theta}}(\hat{\theta}_1, \hat{\theta}_2) & \ldots & \mathrm{cov}_{\underline{\theta}}(\hat{\theta}_1, \hat{\theta}_p) \\
\mathrm{cov}_{\underline{\theta}}(\hat{\theta}_2, \hat{\theta}_1) & \mathrm{var}_{\underline{\theta}}(\hat{\theta}_2) & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
\mathrm{cov}_{\underline{\theta}}(\hat{\theta}_p, \hat{\theta}_1) & \cdots & \cdots & \mathrm{var}_{\underline{\theta}}(\hat{\theta}_p)
\end{bmatrix}.
\end{aligned}$$

This matrix is often referred to as the variance-covariance matrix.

In many cases, only the diagonal entries of the estimator covariance matrix, i.e. the component estimator variances, will be of interest. However, as we will soon see, the entire estimator covariance matrix is very useful for generalizing the scalar parameter CRB.

We can also define the estimator concentration:

$$\begin{aligned}
P_{\underline{\theta}}(\|\hat{\underline{\theta}} - \underline{\theta}\| > \epsilon) &= \int_{\|\hat{\underline{\theta}} - \underline{\theta}\| > \epsilon} f(\hat{\theta}; \theta) d\hat{\theta} \\
&= \int_{\{\underline{x} : \|\hat{\underline{\theta}}(\underline{x}) - \underline{\theta}\| > \epsilon\}} f(\underline{x}; \theta) d\underline{x}
\end{aligned}$$

The first order of business is to extend the CRB to vector parameters, called the *matrix CRB*.

4.5.1 MATRIX CRAMÈR-RAO BOUND (CRB) ON COVARIANCE MATRIX

Let $\underline{\theta} \in \Theta$ be a $p \times 1$ vector and assume:

1. Θ is an open subset of $\mathbb{R}^p$
2. $f(\underline{x}; \underline{\theta})$ is smooth [29] and differentiable in $\underline{\theta}$
3. $\mathrm{cov}_{\underline{\theta}}(\hat{\underline{\theta}})$ and $\mathbf{F}(\underline{\theta})$ (defined below) are non-singular matrices

The matrix CRB for vector valued parameters is the following. For any *unbiased* estimator $\hat{\underline{\theta}}$ of $\underline{\theta}$

$$\mathrm{cov}_{\underline{\theta}}(\hat{\underline{\theta}}) \geq \mathbf{F}^{-1}(\underline{\theta}), \tag{47}$$

where "=" is attained iff the following is satisfied for some non-random matrix $\mathbf{K}_{\underline{\theta}}$

$$\mathbf{K}_{\underline{\theta}} \nabla_{\underline{\theta}} \ln f(\underline{X}; \underline{\theta}) = \hat{\underline{\theta}} - \underline{\theta}. \tag{48}$$

In the case that this tightness condition (48) is satisfied $\hat{\underline{\theta}}$ is said to be an *efficient vector estimator*.

In the matrix CRB (47) $\mathbf{F}(\underline{\theta})$ is the Fisher information matrix, which takes either of two equivalent forms,

$$\begin{aligned} \mathbf{F}(\underline{\theta}) &= E\left[\left(\nabla_{\underline{\theta}} \ln f(\underline{X}; \underline{\theta})\right)\left(\nabla_{\underline{\theta}} \ln f(\underline{X}; \underline{\theta})\right)^T\right] \\ &= -E\left[\nabla^2_{\underline{\theta}} \ln f(\underline{X}; \underline{\theta})\right]. \end{aligned}$$

where we have defined the gradient operator

$$\nabla_{\underline{\theta}} = \left[\frac{\partial}{\partial \theta_1}, \ldots, \frac{\partial}{\partial \theta_p}\right]^T,$$

and the symmetric Hessian (curvature) operator

$$\nabla^2_{\underline{\theta}} = \begin{bmatrix} \frac{\partial^2}{\partial \theta_1^2} & \frac{\partial^2}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2}{\partial \theta_1 \partial \theta_p} \\ \frac{\partial^2}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2}{\partial \theta_2^2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial^2}{\partial \theta_p \partial \theta_1} & \cdots & \cdots & \frac{\partial^2}{\partial \theta_p^2} \end{bmatrix}.$$

The matrix CR Bound (47) has a few more properties than the scalar CRB.

Property 1: The inequality in the matrix bound should be interpreted in the sense of positive definiteness. Specifically if $\mathbf{A}, \mathbf{B}$ are $p \times p$ matrices

$$\mathbf{A} \geq \mathbf{B} \iff \mathbf{A} - \mathbf{B} \geq 0,$$

where $\mathbf{A} - \mathbf{B} \geq 0$ means $\mathbf{A} - \mathbf{B}$ is non-negative definite. This means that, in particular,

$$\underline{z}^T(\mathbf{A} - \mathbf{B})\underline{z} \geq 0$$

for any vector $\underline{z} \in \mathbb{R}^p$, and all eigenvalues of $\mathbf{A} - \mathbf{B}$ are non-negative. For example, choosing $\underline{z} = [1, 0, \ldots, 0]^T$: and $\underline{z} = [1, \ldots, 1]^T$, respectively, $\mathbf{A} \geq \mathbf{B}$, $\mathbf{A} \geq \mathbf{B}$ implies both

$$a_{ii} \geq b_{ii}, \quad \text{and} \quad \sum_{i,j} a_{ij} \geq \sum_{ij} b_{ij}.$$

However, $\mathbf{A} \geq \mathbf{B}$ does NOT mean $a_{ij} \geq b_{ij}$ in general. A simple counterexample is constructed as follows. Let $0 < \rho < 1$ and consider

$$\underbrace{\left[\begin{array}{cc} 2 & 0 \\ 0 & 2 \end{array}\right]}_{\mathbf{A}} - \underbrace{\left[\begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array}\right]}_{\mathbf{B}} = \left[\begin{array}{cc} 1 & -\rho \\ -\rho & 1 \end{array}\right],$$

which has two eigenvalues $1-\rho > 0$ and $1+\rho > 0$. Hence $\mathbf{A}-\mathbf{B} > 0$ while clearly $a_{12} = 0 \not> \rho$.

Property 2: The matrix inequality (47) implies a scalar CRB on the variance of the i-th component of an unbiased vector estimator $\hat{\underline{\theta}}$

$$\mathrm{var}_\theta(\hat{\theta}_i) \geq [\mathbf{F}^{-1}(\underline{\theta})]_{ii},$$

where the right hand side (RHS) denotes the i-th element along the diagonal of the inverse Fisher information matrix.

Property 3. Fisher information matrix is a measure of the average curvature profile of the log likelihood near $\underline{\theta}$

Property 4. Let $\mathbf{F}_n(\underline{\theta})$ be the Fisher information for a sample of n i.i.d. measurements $X_1, \ldots, X_n$. Then, as in the scalar parameter case,

$$\mathbf{F}_n(\underline{\theta}) = n\mathbf{F}_1(\underline{\theta}).$$

Hence $\mathrm{var}_{\underline{\theta}}(\hat{\underline{\theta}}) = O(1/n)$ is also expected for good estimators of multiple unknown continuous valued parameters.

Property 5. Efficient vector estimators only exist for multiparameter exponential families with mean value parameterization

$$f(x;\underline{\theta}) = a(\underline{\theta})b(x)e^{-[\underline{c}(\underline{\theta})]^T[\underline{t}(x)]}$$

and

$$E_{\underline{\theta}}[\underline{t}(X)] = \underline{\theta}.$$

Furthermore, in this case $E[n^{-1}\sum_{i=1}^n \underline{t}(X_i)] = \theta$, $\hat{\underline{\theta}} = n^{-1}\sum_{i=1}^n \underline{t}(X_i)$ is an unbiased efficient estimator of $\underline{\theta}$.

Property 6. If an estimator $\hat{\underline{\theta}}$ satisfies

$$\nabla_{\underline{\theta}} \ln f = \mathbf{K}_{\underline{\theta}}(\hat{\underline{\theta}} - \underline{\theta}),$$

for some non-random matrix $\mathbf{K}_{\underline{\theta}}$ then we can immediately conclude:

1. $\hat{\underline{\theta}}$ is unbiased since, as shown in proof of the multiple parameter CRB;

$$E_{\underline{\theta}}[\nabla_{\underline{\theta}} \ln f(\underline{X};\underline{\theta})] = 0,$$

2. $\hat{\underline{\theta}}$ is efficient and thus its components are UMVU estimators;
3. The covariance of $\hat{\underline{\theta}}$ is given by the inverse Fisher information $\mathbf{F}(\underline{\theta})$;
4. $\mathbf{K}_{\underline{\theta}}$ is the Fisher information $\mathbf{F}(\underline{\theta})$ since

$$E_{\underline{\theta}}[\nabla^2_{\underline{\theta}} \ln f(X,\underline{\theta})] = E_{\underline{\theta}}[\nabla^T_{\underline{\theta}}\nabla_{\underline{\theta}} \ln f(X,\underline{\theta})] = E_{\underline{\theta}}[\nabla_{\underline{\theta}}\{\mathbf{K}_{\underline{\theta}}(\hat{\underline{\theta}} - \underline{\theta})\}]$$

and, by the chain rule and the unbiasedness of $\hat{\underline{\theta}}$

$$E_{\underline{\theta}}[\nabla_{\underline{\theta}}\{\mathbf{K}_{\underline{\theta}}(\hat{\underline{\theta}} - \underline{\theta})\}] = \nabla_{\underline{\theta}}\{\mathbf{K}_{\underline{\theta}}\}E_{\underline{\theta}}[(\hat{\underline{\theta}} - \underline{\theta})\}] + \mathbf{K}_{\underline{\theta}}E_{\underline{\theta}}[\nabla_{\underline{\theta}}\{(\hat{\underline{\theta}} - \underline{\theta})\}] = -\mathbf{K}_{\underline{\theta}}$$

5. The estimator covariance is
$$\mathrm{cov}_{\underline{\theta}}(\hat{\underline{\theta}}) = \mathbf{K}_{\underline{\theta}}^{-1}.$$

Proof of Matrix CR bound:

There are 3 steps in our derivation, which, with one exception, is a direct generalization of the proof of the scalar CRB: (1) show that the gradient of the log-likelihood is zero mean; (2) the correlation between the gradient of the log-likelihood and estimator is constant; (3) the covariance matrix of the concatenated gradient and estimator error gives a relation between Fisher info and estimator covariance.

Step 1. Show $E_{\underline{\theta}}\left[\nabla_{\underline{\theta}} \ln f(\underline{X};\underline{\theta})\right] = 0$.

$$\begin{aligned}
\Rightarrow \quad &= E_{\underline{\theta}}\left[\frac{1}{f(\underline{X};\underline{\theta})}\, \nabla_{\underline{\theta}} f(\underline{X};\underline{\theta})\right] = \int_{\mathcal{X}} \nabla_{\underline{\theta}} f(\underline{x};\underline{\theta}) d\underline{x} \\
&= \nabla_{\underline{\theta}} \underbrace{\int_{\mathcal{X}} f(\underline{x};\underline{\theta}) d\underline{x}}_{=1} = 0.
\end{aligned}$$

Step 2. $E_{\underline{\theta}}\left[\nabla_{\underline{\theta}} \ln f(\underline{X};\underline{\theta})\ (\hat{\underline{\theta}} - \underline{\theta})^T\right] = \mathbf{I}$.

First observe

$$\begin{aligned}
E_{\underline{\theta}}\left[\nabla_{\underline{\theta}} \ln f(\underline{X};\underline{\theta})\ \hat{\underline{\theta}}^T\right] &= E_{\underline{\theta}}\left[\frac{1}{f(\underline{X};\underline{\theta})}\ \nabla_{\underline{\theta}} f(\underline{X};\underline{\theta}) \hat{\underline{\theta}}^T\right] \\
&= \int_{\mathcal{X}} \nabla_{\underline{\theta}} f(\underline{x};\underline{\theta}) \hat{\underline{\theta}}^T(\underline{x}) d\underline{x} \\
&= \nabla_{\underline{\theta}} \underbrace{\int_{\mathcal{X}} f(\underline{x};\underline{\theta}) \hat{\underline{\theta}}^T(\underline{x}) d\underline{x}}_{E_{\underline{\theta}}[\hat{\underline{\theta}}^T] = \underline{\theta}^T} \\
&= \mathbf{I}.
\end{aligned}$$

Now putting this together with result of the previous step

$$\begin{aligned}
&E_{\underline{\theta}}\left[\nabla_{\underline{\theta}} \ln f(\underline{X};\underline{\theta})\ (\hat{\underline{\theta}} - \underline{\theta})^T\right] \\
&= \underbrace{E_{\underline{\theta}}\left[\nabla_{\underline{\theta}} \ln f(\underline{X};\underline{\theta})\ \hat{\underline{\theta}}^T\right]}_{=\mathbf{I}} - \underbrace{E_{\underline{\theta}}\left[\nabla_{\underline{\theta}} \ln f(\underline{X};\underline{\theta})\right]}_{=\underline{0}}\ \underline{\theta}^T.
\end{aligned}$$

Step 3. Define a $2p \times 1$ random vector U:

$$U = \begin{bmatrix} \hat{\underline{\theta}} - \underline{\theta} \\ \nabla_{\underline{\theta}} \ln f(\underline{X};\underline{\theta}) \end{bmatrix}. \tag{49}$$

Since any matrix expressed as an outer product of two vectors is non-negative definite

$$E_{\underline{\theta}}\left[U U^T\right] \geq 0.$$

Using the results of steps 1 and 2, we have

$$E_{\underline{\theta}}\left[UU^T\right] = \left[\begin{array}{cc} \text{cov}_{\underline{\theta}}(\hat{\underline{\theta}}) & \mathbf{I} \\ \mathbf{I} & \mathbf{F}(\underline{\theta}) \end{array}\right] \geq 0.$$

It only remains to apply the result of Sec. 2.4 to the above partitioned matrix to see that this implies that

$$\text{cov}_{\underline{\theta}}(\hat{\underline{\theta}}) - \mathbf{F}^{-1}(\underline{\theta}) \geq 0.$$

An alternative, and more direct, way to show this is to let $\underline{w}$ and $\underline{y}$ be arbitrary p-vectors and define $\underline{v} = \left[\begin{array}{c} \underline{w} \\ \underline{y} \end{array}\right]$. Then, as $\underline{v}^T\ E_{\underline{\theta}}\left[UU^T\right]\ \underline{v} \geq 0$,

$$\underline{w}^T \text{cov}_{\underline{\theta}}(\hat{\underline{\theta}})\underline{w} + 2\underline{w}^T\underline{y} + \underline{y}^T\mathbf{F}(\underline{\theta})\underline{y} \geq 0.$$

Taking $\underline{y} = -\mathbf{F}^{-1}(\underline{\theta})\ \underline{w}$ in the above we obtain

$$\underline{w}^T[\text{cov}_{\underline{\theta}}(\hat{\underline{\theta}}) - \mathbf{F}^{-1}(\underline{\theta})]\underline{w} \geq 0.$$

It remains to obtain the tightness condition ensuring equality in the CRB. Note first that if $\text{cov}_\theta(\hat{\theta}) = \mathbf{F}^{-1}$ then $E_\theta[UU^T]$ necessarily has rank p (see exercises at end of chapter). This can only happen if the random vector U (49) has p linearly independent components. As $\text{cov}_{\underline{\theta}}(\underline{\theta})$ and $\mathbf{F}(\underline{\theta})$ have been assumed non-singular, $\hat{\underline{\theta}} - \underline{\theta}$ can have no linear dependencies and neither does $\nabla_{\underline{\theta}} \ln f$. Hence it can only be that

$$\mathbf{K}_{\underline{\theta}}\nabla_{\underline{\theta}} \ln f = \hat{\underline{\theta}} - \underline{\theta}$$

for some non-random matrix $\mathbf{K}_{\underline{\theta}}$. In other words the gradient of the log likelihood lies in the span of the estimator errors. $\diamond$

We move on to generalizations of MOM and ML estimators to the vector parameter case.

4.5.2 METHODS OF MOMENTS (MOM) VECTOR ESTIMATION

Let $m_k = m_k(\underline{\theta})$ be the k-th order moment of $f(x;\underline{\theta})$. The vector MOM estimation procedure involves finding K moments such that the vector function of $\underline{\theta} \in \mathbb{R}^p$

$$\underline{g}(\underline{\theta}) = [m_1(\theta), \ldots, m_K(\underline{\theta})]$$

can be inverted, i.e., there exists a unique value $\underline{\theta}$ satisfying

$$\underline{\theta} = \underline{g}^{-1}(m_1, \ldots, m_K).$$

As in the scalar case, the MOM estimator is constructed by replacing m_k with its empirical estimate

$$\hat{\underline{\theta}} = \underline{g}^{-1}(\hat{m}_1, \ldots, \hat{m}_K),$$

where $\hat{m}_k = \frac{1}{n}\sum_{i=1}^n X_i^k$.

4.5.3 MAXIMUM LIKELIHOOD (ML) VECTOR ESTIMATION

The vector MLE is an obvious generalization of the scalar MLE

$$\hat{\theta} = \text{argmax}_{\underline{\theta} \in \Theta} f(\underline{X}; \underline{\theta}).$$

For smooth likelihood functions, vector MLEs have several key properties ([29]):

1. Vector MLE's are asymptotically unbiased;

2. Vector MLE's are consistent;

3. Vector MLE's are invariant to arbitrary vector transformations;

$$\underline{\varphi} = \underline{g}(\underline{\theta}) \quad \Rightarrow \quad \hat{\underline{\varphi}} = \underline{g}(\hat{\underline{\theta}});$$

4: Vector MLE's are asymptotically efficient and thus their component estimators are asymptotically UMVU;

5. Vector MLE's are asymptotically Gaussian in the sense

$$\sqrt{n}(\hat{\underline{\theta}}_n - \underline{\theta}) \to \underline{z}, \quad (i.d.)$$

where $\underline{z} \sim \mathcal{N}_p(0, \mathbf{F}_1^{-1}(\underline{\theta}))$ and $\mathbf{F}_1(\underline{\theta})$ is the single sample Fisher information matrix

$$\mathbf{F}_1(\underline{\theta}) \quad = \quad -E_{\underline{\theta}}\left[\nabla^2_{\underline{\theta}} \log f(X_1; \underline{\theta})\right].$$

A couple of examples will illustrate these estimators.

Example 20 *Joint estimation of mean and variance in a Gaussian sample*

This is an extension of Example 20 to the case where both the mean and the variance are unknown. Assume an i.i.d. sample $\underline{X} = [X_1, \ldots, X_n]$ of Gaussian r.v.s $X_i \sim \mathcal{N}(\mu, \sigma^2)$. The unknowns are $\underline{\theta} = [\mu, \sigma^2]$.

The log-likelihood function is

$$l(\underline{\theta}) = \ln f(\underline{x}; \underline{\theta}) = -\frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{k=1}^{n} (x_k - \mu)^2 + c. \tag{50}$$

A. MOM approach to estimation:

We know that $m_1 = \mu$, $m_2 = \sigma^2 + \mu^2$ and thus

$$\mu = m_1, \quad \sigma^2 = m_2 - m_1^2.$$

Hence a MOM estimator of $\underline{\theta}$ is:

$$\begin{aligned}
\hat{\underline{\theta}} &= [\hat{\mu}, \hat{\sigma^2}] \\
&= [\hat{m}_1, \hat{m}_2 - \hat{m}_1^2] \\
&= \left[\overline{X}, \overline{X^2} - \overline{X}^2\right] \\
&= \left[\overline{X}, \overline{(X - \overline{X})^2}\right].
\end{aligned}$$

As usual we denote

$$\begin{aligned}
\overline{X} &= n^{-1}\sum_{k=1}^{n} X_k \\
\overline{(X-\overline{X})^2} &= n^{-1}\sum_{k=1}^{n}(X_k-\overline{X})^2 = \frac{n-1}{n}\mathsf{s}^2,
\end{aligned}$$

and

$$\mathsf{s}^2 = (n-1)^{-1}\sum_{k=1}^{n}(X_k-\overline{X})^2$$

is the sample variance.

B. ML approach.

As $l(\underline{\theta})$ (50) is a concave function (verify that $-\nabla^2_{\underline{\theta}} \ln f$ is positive definite) we can use the likelihood equation (stationary point condition) for finding $\underline{\theta} = \hat{\underline{\theta}}$

$$\underline{0} = \nabla_{\underline{\theta}} \ln f(\underline{x};\underline{\theta}) = \left[\begin{array}{c} \frac{1}{\theta_2}\sum_{k=1}^{n}(x_k-\theta_1) \\ \\ \frac{n/2}{\theta_2} - \frac{1}{2\theta_2^2}\sum_{k=1}^{n}(x_k-\theta_1)^2 \end{array}\right].$$

Therefore,

$$\hat{\theta}_1 = \hat{\mu} = \overline{X}, \qquad \hat{\theta}_2 = \hat{\sigma^2} = \frac{n-1}{n}\mathsf{s}^2,$$

so that the MLE and MOM estimators are identical.

Let's consider the performance of the ML/MOM estimator. The bias and covariance are simple enough to compute (recall that in Sec. 3.4 we showed that $(n-1)\mathsf{s}^2/\sigma^2$ is Chi square distributed with $n-1$ degrees of freedom):

$$\underbrace{E_{\underline{\theta}}[\hat{\mu}] = \mu}_{unbiased}, \qquad \underbrace{E_{\underline{\theta}}[\hat{\sigma^2}] = \left(\frac{n-1}{n}\right)\sigma^2}_{biased};$$

$$\text{var}_{\underline{\theta}}(\overline{X}) = \sigma^2/n;$$

and

$$\text{var}_{\underline{\theta}}(\hat{\sigma^2}) = \left(\frac{n-1}{n}\right)^2 \text{var}_{\underline{\theta}}(\mathsf{s}^2) = 2\sigma^4/n\,\left(\frac{n-1}{n}\right).$$

Since the sample mean and sample variance are uncorrelated (recall Sec. 3.4)

$$\text{cov}_{\underline{\theta}}(\hat{\underline{\theta}}) = \left[\begin{array}{cc} \sigma^2/n & 0 \\ 0 & 2\sigma^4/n\,\left(\frac{n-1}{n}\right) \end{array}\right]. \tag{51}$$

Next we compute the Fisher information matrix by taking the expectation of the Hessian $-\nabla^2_{\underline{\theta}} \ln f(\underline{X};\underline{\theta})$

$$\mathbf{F}(\underline{\theta}) = \left[\begin{array}{cc} n/\sigma^2 & 0 \\ 0 & n/(2\sigma^4) \end{array}\right], \tag{52}$$

giving the CR bound

$$\mathrm{cov}_{\underline{\theta}}(\hat{\underline{\theta}}) \geq \begin{bmatrix} \sigma^2/n & 0 \\ 0 & 2\sigma^4/n \end{bmatrix}. \tag{53}$$

Some interesting observations are the following:

Observation 1. MOM and ML estimators derived above have covariances which violate the CR bound (compare the (2,2) elements of matrices (51) and the RHS of (53)). This is not a contradiction since the ML variance estimator is not unbiased!

Observation 2. Consider the bias-corrected estimator of $[\mu, \sigma^2]^T$

$$\hat{\hat{\underline{\theta}}} = [\overline{X}, \mathsf{s}^2]^T.$$

This estimator is unbiased. Now, as $\mathsf{s}^2 = \left(\frac{n}{n-1}\right)\hat{\sigma^2}$

$$\mathrm{var}_{\underline{\theta}}(\mathsf{s}^2) = \left(\frac{n}{n-1}\right)^2 \mathrm{var}_{\underline{\theta}}(\hat{\sigma^2}),$$

$$\mathrm{cov}_{\underline{\theta}}(\hat{\hat{\underline{\theta}}}) = \begin{bmatrix} \sigma^2/n & 0 \\ 0 & 2\sigma^4/n\left(\frac{n}{n-1}\right) \end{bmatrix} \geq \mathbf{F}^{-1}(\underline{\theta}).$$

We conclude that the bias-corrected estimator's covariance no longer violates the CRB. Indeed, $\overline{X}$ is efficient estimator of μ since

$$\mathrm{var}_{\underline{\theta}}(\hat{\mu}) = [\mathbf{F}^{-1}]_{11} = \sigma^2/n.$$

However, s^2 is not an efficient estimator of σ^2 since

$$\mathrm{var}_{\underline{\theta}}(\mathsf{s}^2) > [\mathbf{F}^{-1}]_{22}.$$

Observation 3. as predicted, the MLE is asymptotically efficient as $n \to \infty$.

$$n\mathrm{cov}_{\underline{\theta}}(\hat{\underline{\theta}}) = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4\left(\frac{n-1}{n}\right) \end{bmatrix} \to \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{bmatrix} = \mathbf{F}_1^{-1}(\underline{\theta}).$$

Observation 4. We can also verify that, as predicted, $[\hat{\mu}, \hat{\sigma^2}]$ is asymptotically Gaussian. It suffices to consider the following results:

a) $\hat{\mu}$ and $\hat{\sigma^2}$ are independent r.v.s;

b) $\sqrt{n}(\hat{\mu} - \mu) = \mathcal{N}(0, \sigma^2)$;

c) $\sqrt{n}(\mathsf{s}^2 - \sigma^2) = \sigma^2\sqrt{n}(\chi^2_{n-1}/(n-1) - 1)$;

d) $\chi^2_\nu \sim \mathcal{N}(\nu, 2\nu)$, $\nu \to \infty$.

Observation 5. We can easily manipulate the condition for equality in the CR bound to find an efficient vector estimator (but not of $\underline{\theta}$ as originally specified!):

$$\nabla_{\underline{\theta}} \ln f(\underline{X}; \underline{\theta}) = \mathbf{K}_{\underline{\theta}} \begin{bmatrix} \overline{X} - \mu \\ \overline{X^2} - (\sigma^2 + \mu^2) \end{bmatrix},$$

where

$$\mathbf{K}_{\underline{\theta}} := \left[\begin{array}{cc} n/\sigma^2 & 0 \\ 0 & n/2\sigma^4 \end{array}\right] \left[\begin{array}{cc} 1 & 0 \\ 2\mu & 1 \end{array}\right]^{-1}.$$

As the sample moments are unbiased estimates of the ensemble moments, we conclude that $\overline{X}, \overline{X^2}$ are efficient estimators of the first moment $E[X] = \mu$ and second (non-central) moment $E[X^2] = \sigma^2 + \mu^2$, respectively.

We continue with another example, which requires special treatment due to functional dependencies that exist between parameters.

Example 21 $\underline{N} = [N_1, \ldots, N_p]^T$ *a multinomial random vector*

The multinomial model is a generalization of the binomial model to more than two categories, "0" and "1", of outcome. Let the outcome $\underline{Z}$ of a single trial be one of the p elementary vectors in $\mathbb{R}^p$, $\underline{e}_1 = [1, 0, \ldots, 0]^T, \ldots, \underline{e}_p = [0, 0, \ldots, 1]^T$, with probabilities $\theta_1, \ldots, \theta_p$, respectively. The vector $\underline{e}_k$ could be a tag attached to the event that a random throw of a die resulted in a face with k dots ($p = 6$) or that a symbol received by a teletype (who remembers those?) corresponds to the k-th letter of the alphabet ($p = 27$). The multinomial model describes the distribution of the sum

$$\underline{N} = [N_1, \ldots, N_p]^T = \sum_{i=1}^n \underline{Z}_i$$

of these vectors obtained after n i.i.d. trials.

The probability of a particular multinomial outcome $\underline{N}$ gives the probability mass function

$$p(\underline{N}; \underline{\theta}) = \frac{n!}{N_1! \cdots N_p!} \theta_1^{N_1} \ldots \theta_p^{N_p}.$$

where $N_i \geq 0$ are integers satisfying $\sum_{i=1}^p N_i = n$ and $\theta_i \in [0, 1]$ are cell probabilities satisfying $\sum_{i=1}^p \theta_i = 1$.

A MOM estimator of $\underline{\theta}$ is obtained by matching the first empirical moment $\underline{N}$ to the first ensemble moment $E_\theta[N] = \underline{\theta} n$. This yields the estimator $\hat{\underline{\theta}} = \underline{N}/n$, or more explicitly

$$\hat{\underline{\theta}} = \left[\frac{N_1}{n}, \ldots, \frac{N_p}{n}\right]$$

To find the MLE of $\underline{\theta}$ we need to proceed with caution. The p parameters $\underline{\theta}$ live in a $p-1$ subspace of $\mathbb{R}^p$ due to total cell probability constraint. We can find the MLE either by reparameterization of the problem or by using Lagrange multipliers. The Lagrange multiplier method will be adopted here.

To account for the constraint we replace the log-likelihood function with the penalized log-likelihood function

$$J(\underline{\theta}) = \ln f(\underline{N}; \underline{\theta}) - \lambda \left(\sum_{i=1}^p \theta_i - 1\right),$$

where λ is a Lagrange multiplier which will be selected. in order to satisfy the constraint.

Now as J is smooth and concave we set the gradient of $J(\underline{\theta})$ to zero to find the MLE:

$$\begin{aligned} 0 = \nabla_{\underline{\theta}} J(\underline{\theta}) &= \nabla_{\underline{\theta}} \left[\sum_{i=1}^{p} N_i \ln \theta_i - \lambda \theta_i \right] \\ &= \left[\frac{N_1}{\theta_1} - \lambda, \ldots, \frac{N_p}{\theta_p} - \lambda \right]. \end{aligned}$$

Thus

$$\hat{\theta}_i = N_i/\lambda, \qquad i = 1, \ldots, p$$

Finally, we find λ by forcing $\underline{\hat{\theta}}$ to satisfy constraint

$$\sum_{i=1}^{p} N_i/\lambda = 1 \quad \Rightarrow \quad \lambda = \sum_{i=1}^{p} N_i = n.$$

The solution to this equation gives the MLE and it is identical to the MOM estimator.

To derive the CRB requires more advanced theory of constrained CR bounds [20] since the θ_i's are linearly dependent.

4.6 HANDLING NUISANCE PARAMETERS

In many cases only a single parameter θ_1 is of direct interest while the other unknowns $\theta_2, \ldots, \theta_p$ are nuisance parameters which are not of interest. For example, in the Gaussian example with both unknown mean and variance, Example 20, the variance may not be of intrinsic interest. In this example, we found that the estimator covariance is diagonal, which implies that there is no correlation between the mean parameter estimation errors and the variance parameter estimation errors. As we will see below, this means that the variance is a rather benign nuisance parameter since knowledge or lack of knowledge of the variance does not affect the variance of the ML mean estimator. We divide the discussion of nuisance parameters into the cases of random and non-random parameters.

CASE I: HANDLING RANDOM NUISANCE PARAMETERS:

For random parameters the average loss only penalizes $\hat{\theta}_1$'s estimation errors:

$$E[c(\hat{\theta}_1, \theta_1)] = \int_{\Theta_1} d\theta_1 \int_{\mathcal{X}} d\underline{x} \; c(\hat{\theta}_1(\underline{x}), \theta_1) f(\underline{x}|\theta_1) f(\theta_1).$$

The prior on θ_1 is computed from the prior on $\underline{\theta}$

$$f(\theta_1) = \int d\theta_2 \ldots \int d\theta_p \; f(\theta_1, \theta_2, \ldots, \theta_p).$$

The conditional density of $\underline{X}$ given θ_1 is therefore

$$f(\underline{x}|\theta_1) = \int d\theta_2 \ldots \int d\theta_p \; f(\underline{x}|\theta_1, \theta_2, \ldots, \theta_p) f(\theta_2, \ldots, \theta_p|\theta_1),$$

yielding the posterior on θ_1

$$f(\theta_1|x) = \int d\theta_2 \ldots \int d\theta_p \; f(\theta_1, \ldots, \theta_p|x).$$

Observe that explicit estimates of $\theta_2, \ldots, \theta_p$ are not required to implement the posterior distribution of θ_1. However, integration (marginalization) of the conditional density over $\theta_2, \ldots, \theta_p$ is required and this may be quite difficult.

CASE II: HANDLING NON-RANDOM NUISANCE PARAMETERS:

The case of non-random parameters is quite different. The average loss still only penalizes for $\hat{\theta}_1$ estimation errors but nonetheless depends on all unknowns:

$$E_{\underline{\theta}}[C] = \int_{\mathcal{X}} c(\hat{\theta}_1(\underline{x}), \theta_1) f(\underline{x}; \underline{\theta})\, d\underline{x}.$$

The maximum Likelihood Estimator of θ_1 is simply

$$\hat{\theta}_1 = \text{argmax}_{\theta_1} \left(\max_{\theta_2, \ldots, \theta_p} \log f(\underline{X}|\theta_1, \theta_2, \ldots, \theta_p) \right).$$

Note that now we require maximization over all nuisance parameters or, equivalently, explicit estimates of the nuisance parameters are necessary.

CR BOUND PREDICTIONS FOR NON-RANDOM NUISANCE PARAMETERS

As above let's say we are interested in unbiased estimation of only the first entry θ_1 in the vector of unknown parameters $\underline{\theta}$. Our derivation of the matrix CRB (47) made the explicit assumption that there existed unbiased estimators of all of the parameters. It turns out that this restriction is unnecessary when only θ_1 is of interest (see exercises).

Assume that $\underline{\theta} = [\theta_1, \ldots, \theta_p]^T$ is an unknown parameter vector. The variance of any unbiased estimator $\hat{\theta}_1$ of θ_1 obeys the lower bound:

$$\text{var}_{\underline{\theta}}(\hat{\underline{\theta}}) \quad \geq \quad [[\mathbf{F}^{-1}(\underline{\theta})]]_{11}, \tag{54}$$

where equality occurs iff there exists a nonrandom vector $\underline{h}_{\underline{\theta}}$ such that

$$\underline{h}_{\underline{\theta}}^T \nabla_{\underline{\theta}} \ln f(\underline{X}; \underline{\theta}) = (\hat{\theta}_1 - \underline{\theta}_1).$$

In (54) $[[\mathbf{A}]]_{ij}$ denotes the ij entry of matrix $\mathbf{A}$, and as before

$$\mathbf{F}(\underline{\theta}) = -E \begin{bmatrix} \frac{\partial^2 l(\underline{\theta})}{\partial \theta_1^2} & \frac{\partial^2 l(\underline{\theta})}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 l(\underline{\theta})}{\partial \theta_1 \partial \theta_p} \\ \frac{\partial^2 l(\underline{\theta})}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 l(\underline{\theta})}{\partial \theta_2^2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial^2 l(\underline{\theta})}{\partial \theta_p \partial \theta_1} & \cdots & \cdots & \frac{\partial^2 l(\underline{\theta})}{\partial \theta_p^2} \end{bmatrix},$$

and $l(\underline{\theta}) = \ln f(\underline{x}; \underline{\theta})$.

Let the Fisher matrix be partitioned as

$$\mathbf{F}(\underline{\theta}) = \begin{bmatrix} a & \underline{b}^T \\ \underline{b} & \mathbf{C} \end{bmatrix},$$

where

* $a = -E_{\underline{\theta}}[\partial^2 \ln f(\underline{X}; \underline{\theta}) / \partial \theta_1^2]$ = Fisher info for θ_1 without nuisance parameters,

* $\underline{b} = -E_{\underline{\theta}}[\partial \nabla_{\theta_2,\ldots,\theta_p} \ln f(\underline{X};\underline{\theta})/\partial\theta_1]$ = Fisher coupling of θ_1 to nuisance parameters,
* $\mathbf{C} = -E_{\underline{\theta}}[\nabla^2_{\theta_2,\ldots,\theta_p} \ln f(\underline{X};\underline{\theta})]$ = Fisher info for nuisance parameters.

Using the partitioned matrix inverse identity (2) the RHS of CRB (54) can be expressed as

$$[[\mathbf{F}^{-1}(\underline{\theta})]]_{11} = \frac{1}{a - \underline{b}^T \mathbf{C}^{-1} \underline{b}}.$$

This gives several insights:

Observation 1: $[[\mathbf{F}^{-1}(\underline{\theta})]]_{11} \geq 1/a = 1/[[\mathbf{F}(\underline{\theta})]]_{11}$. Thus presence of nuisance parameters can only degrade estimator performance;

Observation 2: the amount of degradation is directly proportional to the amount of information coupling between θ_1 and $\theta_2, \ldots, \theta_p$;

Observation 3: no degradation occurs when the Fisher matrix is block diagonal;

4.7 BACKGROUND REFERENCES

One of my favorite introductory texts covering estimation theory is the book on mathematical statistics by Mood, Graybill and Boes [48], mentioned before, which is concise, easy to read, and has many interesting examples and exercises. Nice books on this subject that focus on the Bayesian point of view are Ferguson and [16] and DeGroot [14]. A good survey of Bayesian tools for statistical inference, and estimation in particular, is the book by Tanner [71]. Texts which have more of an engineering flavor are the now classic book by Van Trees [73], and the more recent books by Kay [36], Srinath, Rajasekaran and Viswanathan [67], and Scharf [60]. For a more advanced treatment, requiring some background in real analysis, I like Bickel and Doksum [7], Lehmann [40], and Ibragimov and Has'minskii [29], and Poor [55].

4.8 EXERCISES

4.1 Prove the formula $|a+\Delta| = |a| + \operatorname{sgn}(a)\Delta + [\operatorname{sgn}(a+\Delta) - \operatorname{sgn}(a)](a+\Delta)$ in Sec. 4.2.2.

4.2 Show the equivalence of the two expressions (29) and (30).

4.3 Let $\underline{X} = [X_1, \ldots, X_n]^T$ be a vector of i.i.d. r.v.s X_i which are uniformly distributed over the interval (θ_1, θ_2), $\theta_1 < \theta_2$. Find the maximum likelihood estimator of $\underline{\theta}$.

4.4 Let Z_i, $i = 1, \ldots, n$, be a set of i.i.d. random variables each with the *alpha density*

$$p_\theta(z) = \frac{\beta}{\sqrt{2\pi}\Phi(\alpha) z^2} \exp\left(-\tfrac{1}{2}\left[\alpha - \beta/z\right]^2\right),$$

where $\beta > 0$ is unknown, α is known and $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$ is the standard normal CDF. Assuming that $\alpha = 0$ and that β has an exponential prior density: $p(\beta) = \frac{1}{\sigma_\beta} e^{-\beta/\sigma_\beta}$, where $\sigma_\beta > 0$ is known. Find an expression for the MAP estimate of β. What does the MAP estimate reduce to as $\sigma_\beta \to \infty$ (least informative prior)?

4.5 Let W_i, $i = 1, \ldots, n$, be a set of zero mean i.i.d. Gaussian random variables with variance σ_w^2. Let a be a zero mean Gaussian random variable with variance σ_a^2 which is independent of W_i. The objective is to estimate the value of a given the observation

$$X_i = a + W_i, \quad i = 1, \ldots, n$$

(a) Find the MMSE estimator of a. How does this estimator compare to the MAP and MMAE estimators of a?

(b) Compute the MSE of the MMSE estimator (Hint: express error as a sum of two independent r.v.'s to simplify algebra). What happens to the MSE as $n \to \infty$ or as $\text{SNR} = \sigma_a^2/\sigma_w^2 \to \infty$?

4.6 Let $\underline{X} = [X_1, \ldots, X_n]^T$ be a vector of i.i.d. Gaussian r.v.s with mean μ and variance $\sigma^2 = \mu^2$ ($X_i \sim \mathcal{N}(\mu, \mu^2)$).

(a) Find a method of moments (MOM) estimator of μ based on the first moment.

(b) Find the maximum likelihood estimate of μ.

4.7 Let X_i, $i = 1, \ldots, n$, be an i.i.d. sample from the shifted exponential density $f(x;\theta) = e^{-(x-\theta)}$, $x \geq \theta$, where θ is an unknown parameter $-\infty < \theta < \infty$. Assume that $n > 1$.

(a) Find a MOM estimator of θ.

(b) Find the ML estimator of θ.

(c) Assuming the exponential prior for θ, $f(\theta) = e^{-\theta}$, $\theta \geq 0$, find the MAP estimator, the MMSE estimator, and the MMAE estimator of θ given the i.i.d. sample (be careful with your limits of integration in computing $f(\theta|\underline{x})$!). What happens to these estimators as $n \to \infty$?

(d) Calculate the MSE of each of the estimators derived in part (c) (assume large n). Verify that the MMSE estimator has the lowest MSE.

4.8 The mean square error of a certain unbiased estimator $\hat{\theta}(x)$ of the mean of a measured random variable is equal to $\sigma^2/2$ where $\sigma^2 = \text{var}(x)$. What if anything does this tell you about the distribution of x (Hint: what does the CR bound say about distributions that are impossible)?

4.9 Available are n i.i.d. samples of a random variable X with density

$$f(x;\theta) = \tfrac{1}{2}\frac{1 + 3\theta x^2}{1 + \theta}$$

where $-1 \leq x \leq 1$ and $\theta \geq 0$.

(a) Is this density in the exponential family?

(b) Is the sample mean a sufficient statistic? If so, prove it for general n. If not, give a counterexample, e.g. specialize to $n = 2$.

(c) Find a MOM estimator of θ.

(d) Find the CR bound on estimator variance for any unbiased estimator of θ.

(e) Using either numerical integration (MATLAB) or analysis find the bias and variance of the MOM estimator and compare to the CR bound for large n (e.g. $n = 100$).

4.10 Let the observation X have conditionally uniform density

$$f(x|\theta) = \begin{cases} \frac{1}{\theta}, & 0 < x \leq \theta \\ 0, & o.w. \end{cases}$$

where θ is a random variable with density

$$f_\theta(\theta) = \begin{cases} \theta \exp(-\theta), & \theta \geq 0 \\ 0, & o.w. \end{cases}$$

A useful formula ($v \geq 0$): $\int_v^\infty u e^{-u} du = (v+1)e^{-v}$

(a) Find the MAP estimator of θ.

(b) Find the minimum mean squared error estimator of θ.

(c) Find the minimum mean absolute error estimator of θ.

4.11 Let Z be a single observation having density function

$$p_\theta(z) = (2\theta z + 1 - \theta), \qquad 0 \leq z \leq 1$$

where $-1 \leq \theta \leq 1$.

(a) Assuming that θ is a nonrandom parameter, find and plot the maximum likelihood estimator of θ as a function of Z.

(b) Is the ML estimator unbiased? If so does it achieve the CR bound?

(c) Now assume that θ is a random variable with uniform prior density: $p_\theta(\theta) = \frac{1}{2}, \quad \theta \in [-1, 1]$. Find and plot the minimum mean square error estimator of θ as a function of Z.

(d) Compute the conditional bias $E[\hat{\theta}|\theta] - \theta$ and the conditional MSE $E[(\hat{\theta} - \theta)^2|\theta]$ given θ for each of the estimators of part a and c. Plot the two conditional MSE functions obtained and compare the MSE's of the two estimators. Does one estimator perform uniformly better than the other?

4.12 $\underline{X} = [X_1, \ldots, X_n]^T$ is an i.i.d. observation from the Gamma density

$$X_i \sim f(x|\theta) = \frac{1}{\Gamma(\theta)} x^{\theta-1} e^{-x}, \quad x \geq 0$$

where θ is an unknown non-negative parameter and $\Gamma(\theta)$ is the Gamma function. You should note the useful formulae

$$\Gamma(\theta) = \int_0^\infty x^{\theta-1} e^{-x} dx \quad \textit{and} \quad \frac{\Gamma(\theta + k)}{\Gamma(\theta)} = \theta(\theta + 1) \ldots (\theta + k - 1)$$

.

(a) Find the CR bound on unbiased estimators of θ.

(b) Find the first order MOM estimator of θ by matching ensemble mean to sample mean. Is your estimator unbiased? Compute the variance of your estimator.

4.13 In this exercise you will establish that UMVUE's do not always exist. Let Z be a r.v. with probability mass function

$$p_\theta(z) = \begin{cases} \theta, & z = -1 \\ (1-\theta)^2 \theta^z, & z = 0, 1, 2, \ldots \end{cases}$$

where $\theta \in (0, 1)$.

(a) Define the estimator

$$\hat{\theta}_o(z) = \begin{cases} 1, & z = -1 \\ 0, & z = 0, 1, 2, \ldots \end{cases}.$$

Show that $\hat{\theta}_o$ is an unbiased estimator of θ.

(b) Note that any unbiased estimator $\hat{\theta}$ can be expressed in the form $\hat{\theta} = \hat{\theta}_o + U$ where $U = U(Z)$ is a statistic satisfying $E_\theta[U] = 0$ (any U satisfying this condition is called an *ancillary statistic*). Using this condition and the form for the pmf of Z given above, establish that U must be of the form $U(Z) = aZ$ for some non-random constant a (Hint: Z-transform tables may be helpful).

(c) Now find an expression for the variance of an unbiased $\hat{\theta}$ and show that the value a which minimizes the variance is a function of θ. Hence no single unbiased estimator can achieve minimum variance for all $\theta \in (0,1)$ and therefore no UMVUE for θ exists.

(d) Show that a UMVUE for $\phi = (1-\theta)^2$ *does* exist even though a UMVUE for θ does not exist (Hint: define $\hat{\phi}_o(z) = 1$ for $z = 0$ and $\hat{\phi}_o(z) = 0$, otherwise and repeat the steps in part a through c).

4.14 The observation consists of $x_1, \ldots, x_n$ i.i.d. samples where $x_i \sim f(x|\theta)$ and

$$f(x|\theta) = \begin{cases} \frac{1}{\theta} x^{\frac{1}{\theta}-1}, & 0 \le x \le 1 \\ 0, & o.w. \end{cases}$$

where θ, $0 < \theta < \infty$ is an unknown parameter.

(a) Compute the CR bound on unbiased estimators of θ. Is there an estimator that achieves the bound?

(b) Find the maximum likelihood estimator of θ.

(c) Compute the mean and variance of the maximum likelihood estimator. Specify a function $\varphi = g(\theta)$ for which the maximum likelihood estimator of φ is efficient.

(d) From one of your answers to parts a-c you should be able to derive the following formula

$$\int_0^1 u^\beta \ln\left(\frac{1}{u}\right) du = \frac{1}{(1+\beta)^2}, \qquad \beta > -1.$$

4.15 The measurement $\underline{x} = [x_1, \ldots, x_n]^T$ is i.i.d. Gaussian with unknown mean μ and variance σ^2.

(a) Show that the sample mean $\overline{x_i} = n^{-1}\sum_{i=1}^n x_i$ and sample variance $\mathsf{s}^2 = (n-1)^{-1}\sum_{k=1}^n (x_k - \overline{x_i})^2$ are unbiased estimators and that they are uncorrelated *and* independent random variables (**Hint**: show that the Gaussian random variables $x_i - \overline{x_i}$ and $\overline{x_i}$ are uncorrelated for $i = 1, \ldots, n$).

(b) Using the results of part (a) derive the covariance matrix for the estimator $\underline{\hat{\theta}} = [\overline{x_i}, \mathsf{s}^2]^T$. (Hint: to save yourself lots of algebra you should represent $\mathsf{s}^2 = \mathsf{s}^2(\underline{x})$ in terms of σ^2 and the sample variance $\mathsf{s}^2(\underline{z})$ for $\underline{z}$ a vector of n i.i.d. zero mean unit variance Gaussian variables. Then use the representation (ch. 3 of course notes) $\mathsf{s}^2(\underline{z}) = \frac{1}{n-1}\chi_{n-1}$ and properties of the Chi square r.v. to find the expression for variance of s^2).

(c) Derive the CR bound on the covariance matrix of any unbiased estimator $\underline{\hat{\theta}}$ of $\underline{\theta} = [\theta_1, \theta_2]^T = [\mu, \sigma^2]^T$. Compare to the result of part (b).

4.16 Show that if the CR bound is attained with equality then $E_\theta[UU^T]$ has rank p, where U is given by (49). (Hint: show that the matrix

$$E_{\underline{\theta}}\left[UU^T\right] = \begin{bmatrix} \mathbf{F}^{-1}(\underline{\theta}) & \mathbf{I} \\ \mathbf{I} & \mathbf{F}(\underline{\theta}) \end{bmatrix}$$

has rank p.)

4.17 An alternative approach to parameter estimation is called the "quantile matching method" and you will explore this method here. Let $f(x;\theta)$ be a density of the continuous r.v. X parameterized by the scalar parameter θ and define the *theoretical cdf* $F(x;\theta) = \int_{-\infty}^x f(u;\theta)du$.

For n i.i.d. realizations $\{X_i\}_{i=1}^n$ from $f(x;\theta)$ define the *empirical cdf* as the fraction of X_i's which are less than or equal to x:

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^n I_{(-\infty,x]}(X_i)$$

where $I_A(y)$ equals 1 if $y \in A$ and zero otherwise (the indicator function of set A).

(a) Derive the mean $E_\theta[\hat{F}(x)]$ and covariance $\mathrm{cov}_\theta(\hat{F}(x), \hat{F}(y))$ of $\hat{F}$. Show that $\hat{F}(x)$ is an asymptotically consistent estimator of $F(x;\theta)$.

(b) The quantile matching estimate (QME) $\hat{\theta}$ is defined as that value of t which minimizes

$$\int_{-\infty}^{\infty} |F(x;t) - \hat{F}(x)|^2 dx \tag{55}$$

Let θ be a location parameter: $f(x;\theta) = f(x-\theta)$. Using the definition (55), show that $\hat{\theta}$ must satisfy the following equation (Hint: use integration by parts):

$$\int_{-\infty}^{\infty} f(x-\hat{\theta})\hat{F}(x)dx - 1/2 = 0. \tag{56}$$

Show that if $\hat{\theta}$ is the unique solution to (56) it is an asymptotically consistent estimator of θ (Hint: for $\hat{\theta} = t$ fixed and non-random, compute mean square value of left hand side of (56) and show that as $n \to \infty$ it goes to a function of t which equals zero at $t = \theta$).

(c) Using matlab, or other software application of your choice, simulate the QME and the MLE for the following cases:

i. $f(x;\theta)$ Gaussian with variance 1 and mean θ.

ii. $f(x;\theta) = \alpha e^{-\alpha(x-\theta)} I_{[\theta,\infty)}(x)$ (shifted exponential) with $\alpha = 1$.

Run the above simulations 50-100 times each for the cases of $n = 1, 5, 10, 15, 20, 25$ observations, respectively. Using the results of your simulations find and plot as a function of n: 1) the average mean-squared error for MLE and QME estimators; 2) the average quantile squared error (55) evaluated at $t = \hat{\theta}$ (you should show 4 different plots). Also generate a couple of representative plots of the objective function (55) as a function of t for the Gaussian and shifted exponential cases above. Comment on what can be concluded from your simulation study.

4.18 Available are n i.i.d. samples of a discrete random variable X with probability mass function $P(X = k) = p(k;\theta)$, given by

$$p(k;\theta) = \begin{cases} \left(\frac{\theta}{1+\theta}\right)^{k-k_o} \frac{1}{1+\theta}, & k = k_o, k_o+1, \ldots \\ 0, & o.w. \end{cases},$$

where k_o is a known non-negative integer and θ is unknown with $0 \leq \theta < \infty$. (A potentially useful identity: $\sum_{k=0}^{\infty} k a^k = a/(1-a)^2$).

(a) Is this density in the exponential family with mean value parameterization? Find a one dimensional sufficient statistic for θ.

(b) Find a MOM estimator of θ.

(c) Find the ML estimator of θ.

(d) Find the Fisher information on estimator variance for any unbiased estimator of θ. Are either of the estimators of part (b) or part (c) efficient?

4.19 Available is a single measurement of a random variable W. The model for W is

$$W = (1 - Z)X + ZY,$$

where Z is Bernoulli with $P(Z = 0) = P(Z = 1) = 1/2$, X is Gaussian with zero mean and variance σ^2, and Y is Gaussian with mean μ and variance σ^2. Assume that μ and σ^2 are known and that X, Y, Z are independent.

(a) Find the posterior distribution of Z.

(b) Find the minimum mean squared error estimator of Z. Plot the estimator as a function of W.

(c) Find the MAP estimator of Z. Plot the estimator as a function of W.

(d) Find the affine minimum mean squared error estimator of Z. Plot the estimator as a function of W.

4.20 Let $X_1, X_2, \ldots, X_n$ be i.i.d. variables with the *standard Pareto density*:

$$f(x;\theta) = \begin{cases} \theta c^\theta x^{-(\theta+1)}, & x \geq c \\ 0, & o.w. \end{cases}$$

where $c > 0$ is known and $\theta > 0$ is unknown.

(a) Is $f(x;\theta)$ a member of the exponential family? Why or why not?

(b) Find a one dimensional sufficient statistic for θ given $X_1, X_2, \ldots, X_n$.

(c) Find the Fisher information and state the CR bound for unbiased estimators of θ.

(d) Derive the maximum likelihood estimator $\hat{\theta}$ of θ.

(e) Is your estimator efficient?

4.21 Let $X_1, X_2, \ldots, X_n$ be i.i.d. variables with the *generalized Pareto* density:

$$f(x;\theta) = \begin{cases} c\theta^c x^{-(c+1)}, & x \geq \theta \\ 0, & o.w. \end{cases}$$

where $c > 0$ is known and $\theta > 0$ is unknown.

(a) Is $f(x;\theta)$ a member of the exponential family? Why or why not?

(b) Find a one dimensional sufficient statistic for θ given $X_1, X_2, \ldots, X_n$.

(c) Derive the maximum likelihood estimator $\hat{\theta}$ of θ.

4.22 The posterior density of a scalar parameter θ given an observation $\underline{x} = [x_1, \ldots, x_n]^T$ is a function of the form $f(\theta|\underline{x}) = g(\overline{x_i} - \theta)$ where $\overline{x_i}$ is the sample mean and g is an integrable function satisfying $g(-u) = g(u)$ and $g(0) > g(u)$, $u \neq 0$. Derive the MAP, CME and CmE estimators of θ.

4.23 The CRB has several generalizations that we explore in this problem for scalar parameters θ of a density $f_\theta(x)$.

(a) Define the finite difference $\delta f = (f_{\theta+\Delta} - f_\theta)/\Delta$. Show that for any unbiased estimator $\hat{\theta}$ of non-random θ

$$\mathrm{var}_\theta(\hat{\theta}) \geq \frac{1}{E_\theta\left[(\delta f_\theta / f_\theta)^2\right]}$$

with equality iff $\delta f_\theta / f_\theta = k_\theta(\hat{\theta} - \theta)$ for non-random constant k_θ. The above bound is called the Chapman Robbins version of the Barankin bound

(b) Show that the bound of part (a) implies the CRB in the case that θ is a non-random continuous parameter and f_θ is smooth (Hint: take limit as $\Delta \to 0$).

(c) When θ is a random variable with prior density $p(\theta)$ show that

$$E[(\hat{\theta} - \theta)^2] \geq \frac{1}{J}$$

where

$$J = E\left[(\delta p(\theta|X)/p(\theta|X))^2\right]$$

and $\delta p(\theta|X) = (p(\theta + \Delta|X) - p(\theta|X))/\Delta$. Here the expectation E is taken over both X and θ.

4.24 Let $g(x; \phi_1)$ and $h(x; \phi_2)$ be densities where ϕ_1, ϕ_2 are unknown scalar parameters. The arithmetic epsilon mixture model for X is:

$$f_{\mathcal{A}}(x; \theta) = (1 - \epsilon)g(x; \phi_1) + \epsilon h(x; \phi_2)$$

where $0 \leq \epsilon \leq 1$ and $\underline{\theta} = [\phi_1, \phi_2, \epsilon]^T$. The geometric epsilon mixture model for X is:

$$f_{\mathcal{G}}(x; \underline{\theta}) = \frac{1}{d(\underline{\theta})}\, g^{1-\epsilon}(x; \phi_1) h^{\epsilon}(x; \phi_2), \tag{57}$$

where

$$d(\underline{\theta}) = \int g^{1-\epsilon}(x; \phi_1) h^{\epsilon}(x; \phi_2) dx$$

is a normalizing constant (related to the Rènyi ϵ-divergence between g and h). From this exercise you will appreciate that the mixture $f_{\mathcal{G}}$ is easier to deal with than $f_{\mathcal{A}}$ for the purposes of investigating CR bounds, detectors and estimators. Assume that g and h are members of the exponential family of densities.

(a) Show that the three parameter density $f_{\mathcal{G}}(x; \underline{\theta})$ is a member of the exponential family. Show that $f_{\mathcal{A}}(x; \underline{\theta})$ is not a member of this family.

(b) Derive expressions for the six distinct entries of the Fisher information matrix (FIM) for jointly estimating the parameters $\underline{\theta}$ from n i.i.d. observations from $f_{\mathcal{G}}$. An explicit expression for the FIM does not generally exist for the standard mixture model $f_{\mathcal{A}}$.

(c) For n i.i.d. observations from $f_{\mathcal{G}}$ give a condition on the parameter vector $\underline{\theta}$ which guarantees that an efficient estimator exist for $\underline{\theta}$, i.e. for which the inverse FIM is an achievable lower bound on the covariance of unbiased estimators of $\underline{\theta}$ (Hint: what is the mean value parameterization as defined by (28)?).

(d) In the sequel of this exercise we specialize $f_{\mathcal{G}}$ to the case of a geometric mixture of two exponential densities

$$g(x; \theta) = \phi_1 \exp(-x\phi_1), \quad h(x; \theta) = \phi_2 \exp(-x\phi_2), \tag{58}$$

where $x, \phi_1, \phi_2 > 0$. Derive an expression for $d(\underline{\theta})$. Is the CR bound achievable for this model?

(e) Let n i.i.d. realizations be available from the geometric mixture $f_{\mathcal{G}}$ specified by (57) and (58). By evaluating the gradient of the likelihood function, find a set of (non-linear) equations which must be satisfied by the MLE of $\underline{\theta}$. Using these equations, and assuming that ϕ_1, ϕ_2 are known, find an explicit expression for the MLE of ϵ.

4.25 Let S and X be jointly Gaussian distributed with means and variances

$$\begin{aligned}&E[S] = \mu_S,\ E[X] = \mu_X,\\ &\text{var}(S) = \sigma_S^2,\ \text{var}(X) = \sigma_X^2\\ &\text{cov}(S,X) = \rho\,\sigma_S\sigma_X.\end{aligned}$$

Specifically the joint density is bivariate Gaussian

$$f_{S,X}(s,x) = \frac{1}{2\pi\sigma_S\sigma_X\sqrt{1-\rho^2}}\exp\left(\frac{-1}{2(1-\rho^2)}\left[\frac{(s-\mu_S)^2}{\sigma_S^2} - 2\rho\frac{(s-\mu_S)(x-\mu_X)}{\sigma_S\sigma_X} + \frac{(x-\mu_X)^2}{\sigma_X^2}\right]\right).$$

(a) By integrating the joint density over s, show that the marginal density f_X of X is a univariate Gaussian density with mean parameter μ_X and variance parameter σ_X^2.

(b) Using the above to show that the conditional density $f_{S|X}(s|x)$ of S given X is univariate Gaussian with mean and variance parameters

$$\begin{aligned}\mu_{S|X}(x) &= \mu_S + \rho\frac{\sigma_S}{\sigma_X}(x-\mu_X),\\ \sigma_{S|X}^2 &= (1-\rho^2)\sigma_S^2.\end{aligned}$$

Note that while the mean parameter depends on x the variance parameter is independent of x.

(c) Using this form for the conditional density show the mean and variance parameters are precisely the conditional mean and variance of S given $X = x$, respectively.

4.26 A charity box is placed in a mall. The box can only accept quarters. With probability p (a deterministic quantity), a (good) person would come and place a quarter in the box, thus incrementing the number of quarters in the box by one. With probability $1-p$, a (bad) person would come and empty the box, thus setting the number of quarters in the box to zero.

Assuming stationarity, it can be shown that the probability that k quarters will be observed at the end of the d-th day is

$$P(T(d) = k) \;=\; p^k(1-p).$$

(Notation: $T(d)$ is the random variable representing the number of quarters in the box at the end of the d-th day.) In the following you should assume that $T(1), T(2), \ldots,$ are independent identically distribute (i.i.d) random variables.

(a) *Maximum Likelihood and Efficiency:* To estimate the percentage of good people p, the box monitor counts the number of quarters in the box at the end of each day, D days in a row.
- Write down the joint PDF of the vector of number of quarters observed $[T(1), T(2), \ldots, T(D)]$.
- Find the ML estimator of p given $T(1) = k_1$, $T(2) = k_2$, …, $T(D) = k_D$.
- Is the ML estimator $\hat{p}_{ML}$ efficient ?

(b) *Method of Moments:* Define the the average number of quarters observed as $\bar{k} = \frac{1}{D}\sum_{d=1}^{D} k_d$.
- Find the expected value of the average number of quarters observed $E[\bar{k}]$ (hint: $\sum_{n=0}^{\infty} np^n = \frac{p}{(1-p)^2}$).
- Based on this result, suggest a method of moments estimator for p.

(c) *Efficiency and the CRB:* To investigate how well the charity box is doing, a new measure is considered $\gamma = \frac{p}{1-p}$, the ratio of the percentage of good people to the percentage of bad people, otherwise known as the good-to-bad ratio (GBR).

- Is the ML estimator of the GBR $\hat{\gamma}_{ML}$ efficient ?
- Find the ML estimator of the GBR $\hat{\gamma}_{ML}$.
- Find the Cramér-Rao bound (CRB) on the MSE of an unbiased estimator for the GBR.
- Find the MSE of the ML estimator of the GBR.

4.27 Here you will show that the MLE is invariant to arbitrary functional transformations of the parameter. Let θ be a scalar parameter with range $\Theta = (-\infty, \infty)$, assume the sample $\underline{X}$ has j.p.d.f $f(\underline{x}; \theta)$, and that there exists a unique MLE $\hat{\theta}$. Given a transformation g define the new parameter $\varphi = g(\theta)$.

(a) Assume that g is monotone, i.e. $g(\theta)$ is 1-1 invertible over all Θ. Show that the MLE of φ is

$$\hat{\varphi} = g(\hat{\theta}).$$

(b) Next assume that g is smooth in the sense of piecewise monotonicity, i.e., there exists a partition of Θ into intervals $(-\infty, \theta_1], (\theta_1, \theta_2], \ldots, (\theta_M, \infty)$ such that g is monotone over each of these intervals (M may not be finite). Define the integer function h by: $h(\theta) = k$, if θ is in the k-th interval, $k = 1, \ldots, M + 1$. Show that the scalar-to-vector mapping $\theta \to [g(\theta), h(\theta)]$ is 1-1 invertible.

(c) Using result of (b) show that the MLE is invariant to piecewise monotone functional transformation.

4.28 Derive the CR bound (54) on the variance of an unbiased scalar estimator $\hat{\theta}_1$ of θ_1 when the rest of the parameters $\theta_2, \ldots, \theta_p$ in $\underline{\theta}$ are unknown nuisance parameters. Do not assume that the nuisance parameters have unbiased estimators (Hint: define $\underline{U} = [\hat{\theta}_1 - \theta_1, \nabla^T_{\underline{\theta}} \ln f(\underline{X}; \underline{\theta})]^T$ and proceed as in the proof of the matrix CRB).

4.29 A sequence of measurements $X_1, \ldots, X_n$ are i.i.d. with marginal density

$$f_{X_i}(x; \theta) = \frac{\theta}{x^2} e^{-\frac{\theta}{x}}, \quad x > 0$$

where $\theta > 0$ is an unknown parameter.

(a) For part (a) and (b) assume that θ is non-random. Is this density a member of the exponential family? Find a one dimensional sufficient statistic for θ.

(b) Find the maximum likelihood estimator of θ.

(c) For part (c) and (d) assume that θ is a random variable having density

$$f(\theta) = e^{-\theta}, \quad \theta > 0.$$

Find the MAP estimator of θ.

(d) Find the minimum mean squared error estimator of θ and compare to your result in part (c). Hint: $\int_0^\infty \alpha^n e^{-\alpha} d\alpha = n!$.

4.30 Show that the vector conditional mean estimator $\hat{\underline{\theta}}_{CME}$ of a random vector parameter $\underline{\theta}$ satisfies the property that, for any other estimator $\hat{\underline{\theta}}$

$$E[(\underline{\theta} - \hat{\underline{\theta}})(\underline{\theta} - \hat{\underline{\theta}})^T] \geq E[(\underline{\theta} - \hat{\underline{\theta}}_{CME})(\underline{\theta} - \hat{\underline{\theta}}_{CME})^T],$$

where the matrix inequality $\mathbf{A} \geq \mathbf{B}$ is interpreted in terms of non-negative definiteness of $\mathbf{A} - \mathbf{B}$.

4.31 Let θ be a nonrandom vector parameter of some smooth (in θ) density function $f(x;\theta)$. Show that $E_\theta\left[\nabla_\theta \ln f(X;\theta)(\nabla_\theta \ln f(X;\theta))^T\right] = E_\theta[-\nabla^2_\theta \ln f(X;\theta)]$.

4.32 Assume that X is a sample from a density in an exponential family with scalar parameter θ having the mean value parameterization ($E_\theta[t(X)] = \theta$, recall discussion in Sec. 3.5.4). Assuming the Fisher information $F(\theta)$ exists show that

$$F(\theta) \quad = \quad 1/\text{var}_\theta(t(X)). \tag{59}$$

Now show that if one has an i.i.d. sample $\underline{X} = [X_1, \ldots, X_n]^T$ from such a density then $\hat{\theta} = n^{-1}\sum_{i=1}^n t(x_i)$ is an unbiased and efficient estimator of θ.

End of chapter

5 LINEAR ESTIMATION

In the previous chapter we discussed several strategies for estimating parameters given a model $f(x;\theta)$ for the probability distribution of the measurements. These strategies all require precise knowledge of this model which may not always be available. Furthermore, even when one has full confidence in the model these strategies usually yield estimators that are non-linear function of the measurements and whose implementation may be difficult, e.g. involving analytical maximization of a complicated density or analysis of its moments as a function of θ. In this chapter we present an alternative linear estimation approach which only requires knowledge of the first two moments or empirical estimates of these moments. While linear methods do not have the optimality properties of optimal Bayes estimators, such as MAP or CME, they are very attractive due to their simplicity and to their robustness to unknown variations in higher order moments.

Linear estimation theory starts out by assuming random parameters, adopting a squared error loss function, and then seeking to minimize the mean squared error over all estimator functions defined as linear or affine functions of the measurements. It turns out that this linear minimum mean squared error (LMMSE) problem can be recast as minimization of a norm in a linear vector space. This leads to an elegant and intuitive geometric interpretation of the optimal LMMSE estimator via the projection theorem and orthogonality condition of min-norm problems on linear vector spaces. The resultant LMMSE estimator depends on the mean and variance of the measurement, the mean of the parameter, and the covariance of the measurements and parameter. Not surprisingly, when the measurements and parameters are jointly Gaussian distributed the affine estimator is equivalent to the optimal conditional mean estimator. When the means and covariances are not known *a priori* an analogous nonstatistical linear least squares (LLS) estimation theory can be developed, leading to the well known problem of linear regression.

As usual the main ingredients for linear estimation will be the vector of measurements $\underline{x} = [x_1, \ldots, x_n]^T$ and the vector of parameters $\underline{\theta} = [\theta_1, \ldots, \theta_p]^T$. In Sec. 5.1 we cover the case where these vectors are realizations of random variables with known first and second (ensemble) moments. In Sec. 5.6 we turn to the case where these moments are unknown.

5.1 MIN MSE CONSTANT, LINEAR, AND AFFINE ESTIMATION

First we will assume that $\underline{x}$ and $\underline{\theta}$ are realizations of two random vectors $\underline{X}$ and $\underline{\theta}$. Similarly to the last chapter, we use the notation $E[\theta] = \int f(\underline{\theta}) d\underline{\theta}$ to denote expectation. However, in this section we will never refer to the density $f(\underline{\theta})$ explicitly since we will only assume knowledge of its first and second order moments. The overall objective is to find the solution to the minimization

$$\min_{\hat{\underline{\theta}}} \mathrm{MSE}(\hat{\underline{\theta}}) = \min_{\hat{\underline{\theta}}} E[\|\underline{\theta} - \hat{\underline{\theta}}(\underline{X})\|^2]$$

where the expectation is over both $\underline{\theta}$ and $\underline{X}$ and the minimization is restricted to constant, linear or affine functions $\hat{\underline{\theta}}$ of $\underline{X}$. The norm in this minimization is the standard euclidean 2-norm $\|\underline{u}\| = \sqrt{\underline{u}^T \underline{u}}$. We first specialize to scalar parameters to eliminate unnecessary complications in the derivations to follow. We extend the treatment to vector parameters in Sec. 5.5.

5.1.1 BEST CONSTANT ESTIMATOR OF A SCALAR RANDOM PARAMETER

This is the simplest possible estimator structure as the constant estimator $\hat{\theta} = c$ does not depend on the measurements. It turns out that the best constant estimator only depends on the mean of the parameter and no additional information about the measurements or the parameter distributions is needed.

The problem is to find the constant $\hat{\theta} = c$ that minimizes MSE

$$\text{MSE}(c) = E[(\theta - c)^2].$$

Solution: $\hat{\theta} = E[\theta]$ is the best constant estimator.

As the MSE is a quadratic function of c this can easily be proven by setting the derivative $\frac{d}{dc}\text{MSE}(c)$ to zero. Another, more direct way of deriving this solution is add and subtract the mean $E[\theta]$ from $\theta - c$ and expand the square in MSE(c) to obtain a sum of two terms, one of which is zero:

$$\begin{aligned}
\text{MSE}(\hat{\theta}) &= E[((\theta - E[\theta]) - (c - E[\theta]))^2 \\
&= E[(\theta - E[\theta])^2] + (E[\theta] - c)^2 - 2(E[\theta] - c)\underbrace{E[\theta - E[\theta]]}_{=0} \\
&= E[(\theta - E[\theta])^2] + (E[\theta] - c)^2.
\end{aligned}$$

As only the second term in the last line depends on c and it is non-negative it is obvious that $c = E[\theta]$ is the best constant estimator. The resultant min MSE is immediate:

$$\min_c \text{MSE}(c) = E[(\theta - E[\theta])^2],$$

which is just the prior variance var(θ) of θ.

Since the constant estimator uses no information about the measurements we can expect that any good $\underline{X}$-dependent estimator of θ will have lower MSE than var(θ).

5.2 BEST LINEAR ESTIMATOR OF A SCALAR RANDOM PARAMETER

The next step up in complexity is an estimator that depends linearly on $\underline{X}$

$$\hat{\theta} = \underline{h}^T \underline{X},$$

where $\underline{h} = [h_1, \ldots, h_n]^T$ is a set of linear coefficients to be determined. It will be seen that to implement the linear minimum MSE (LMMSE) estimator we require the second moment matrix,

$$\mathbf{M}_X = E[\underline{X}\,\underline{X}^T]$$

and the cross-moment vector

$$\underline{m}_{X,\theta} = E[\underline{X}\theta].$$

We will assume that $\mathbf{M}_X$ is an invertible matrix.

The problem is to find the coefficient vector $\underline{h}$ that minimizes MSE

$$\text{MSE}(\underline{h}) = E[(\theta - \underline{h}^T \underline{X})^2].$$

Solution: $\hat{\theta} = \underline{m}_{X,\theta}^T \mathbf{M}_X^{-1} \, \underline{X}$ is the LMMSE estimator.

To derive this solution we note that the MSE is a quadratic function of $\underline{h}$:

$$\begin{aligned} \mathrm{MSE}(\hat{\theta}) &= E[(\theta - \hat{\theta})^2] = E[(\theta - \underline{h}^T \underline{X})^2] \\ &= \underline{h}^T E[\underline{X}\,\underline{X}^T]\underline{h} + E[(\theta)^2] - \underline{h}^T E[\underline{X}\theta] - E[\theta \underline{X}^T]\underline{h} \end{aligned}$$

The $\underline{h}$ that minimizes this quadratic form can be found by differentiation or by completion of the square. Following the former route (see Sec. 2.4.3 for review of derivatives of functions of a vector variable) we obtain:

$$\begin{aligned} \underline{0}^T &= \nabla_{\underline{h}} \mathrm{MSE}(\underline{h}) = \left[\frac{\partial}{\partial h_1}, \ldots, \frac{\partial}{\partial h_n}\right] \mathrm{MSE}(\hat{\theta}) \\ &= 2\left(\underline{h}^T E[\underline{X}\,\underline{X}^T] - E[\theta \underline{X}^T]\right) \end{aligned}$$

Therefore the optimal $\underline{h}$ satisfies the equation:

$$E[\underline{X}\,\underline{X}^T]\underline{h} = E[\underline{X}\theta]$$

Assuming non-singular $\mathbf{M}_X = E[\underline{X}\,\underline{X}^T]$ this is equivalent to

$$\underline{h} = \mathbf{M}_X^{-1} \underline{m}_{X\theta}$$

and the optimal linear estimator is

$$\hat{\theta} = \underline{m}_{X\theta}^T \mathbf{M}_X^{-1} \, \underline{X},$$

as claimed.

By plugging the optimal solution back into $\mathrm{MSE}(\underline{h})$ it is easy to see that the minimum MSE over linear estimators is

$$\mathrm{MSE}_{\min} = E[\theta^2] - \underline{m}_{X\theta}^T \mathbf{M}_X^{-1} \underline{m}_{X\theta}.$$

Note that, as $\mathbf{M}_X^{-1}$ is positive definite, this MSE can never exceed the *a priori* second moment $E[|\theta|^2]$ of θ. If the parameter is zero mean then $E[\theta^2] = E[(\theta - E[\theta])^2] = \mathrm{var}(\theta)$, i.e., the second moment is equal to the *a priori* variance and the LMMSE estimator generally outperforms the best constant estimator $\hat{\theta} = E[\theta] = 0$. However, if $E[\theta] \neq 0$ then $E[\theta^2] > E[(\theta - E[\theta])^2]$ and the linear estimator may not even do as well as the constant estimator. The problem is that the LMMSE estimator can be a biased estimator of θ in the sense that its average bias $E[\hat{\theta}] - E[\theta] \neq 0$ unless $E[\theta] = 0$. The way to handle this bias is to generalize the class of linear estimators to the class of affine estimators.

5.3 BEST AFFINE ESTIMATOR OF A SCALAR R.V. θ

An affine estimator also depends linearly on $\underline{X}$ but incorporates a constant term to control bias

$$\hat{\theta} = \underline{h}^T \underline{X} + b = \underline{h}^T (\underline{X} - E[\underline{X}]) + c,$$

where $c = b + \underline{h}^T E[\underline{X}]$ and b are just different parameterizations of the bias controlling constant. It will be easier to deal with c here. The objective is to determine the best coefficients $\{\underline{h} =$

$[h_1, \ldots, h_n]^T, c\}$. To implement the affine minimum MSE estimator we require knowledge of the means $E[\underline{X}]$, $E[\theta]$, the (assumed invertible) covariance matrix,

$$\mathbf{R}_X = \text{cov}(\underline{X}) = E[(\underline{X} - E[\underline{X}])(\underline{X} - E[\underline{X}])^T],$$

and the cross-correlation vector

$$\underline{r}_{X,\theta} = \text{cov}(\underline{X}, \theta) = E[(\underline{X} - E[\underline{X}])(\theta - E[\theta])].$$

The problem is to find the vector $\underline{h}$ and the scalar c that minimizes MSE

$$\text{MSE}(\underline{h}, c) = E[(\theta - \underline{h}^T(\underline{X} - E[\underline{X}]) - c)^2].$$

Solution: $\hat{\theta} = E[\underline{\theta}] + \underline{r}_{X,\theta}^T \mathbf{R}_X^{-1}(\underline{X} - E[\underline{X}])$ is the best affine estimator.

To derive this solution we again note that the MSE is a quadratic function of the unknowns $\underline{h}$, c

$$\begin{aligned}
\text{MSE}(\hat{\theta}) &= E[|\theta - \hat{\theta}|^2] = E[|(\theta - c) - \underline{h}^T(\underline{X} - E[\underline{X}])|^2] \\
&= \underline{h}^T \underbrace{E[(\underline{X} - E[\underline{X}])(\underline{X} - E[\underline{X}])^T]}_{\mathbf{R}_X} \underline{h} + E[|\theta - c|^2] \\
&\quad -\underline{h}^T \underbrace{E[(\underline{X} - E[\underline{X}])(\theta - c)]}_{\underline{r}_{X,\theta}} - \underbrace{E[(\theta - c)(\underline{X} - E[\underline{X}])^T]}_{\underline{r}_{\theta,X}} \underline{h} \\
&= \underline{h}^T \mathbf{R}_X \underline{h} + E[|\theta - c|^2] - 2\underline{h}^T \underline{r}_{X,\theta}.
\end{aligned}$$

Note that the only dependence on c is through a single term that is minimized by choosing $c = E[\theta]$. As for $\underline{h}$, the minimizer can be found by differentiation,

$$\underline{0} \quad = \quad \nabla_{\underline{h}} \text{MSE} = \underline{h}^T \mathbf{R}_X - \underline{r}_{\theta,X},$$

which leads to the equation for the minimizer $\underline{h}$

$$\mathbf{R}_X \underline{h} = \underline{r}_{X,\theta}.$$

When $\mathbf{R}_X$ is non-singular this is equivalent to

$$\underline{h} = \mathbf{R}_X^{-1} \underline{r}_{X,\theta}$$

and the optimal affine estimator is therefore

$$\hat{\theta} = E[\theta] + \underline{r}_{X,\theta}^T \mathbf{R}_X^{-1}(\underline{X} - E[\underline{X}]).$$

Unlike the linear estimator, the affine estimator is on-the-average unbiased in the sense that $E[\hat{\theta}] = E[\theta]$.

The minimum MSE attained by this estimator is simply computed as

$$\text{MSE}_{\min} \quad = \quad \text{var}(\theta) - \underline{r}_{X,\theta}^T \mathbf{R}_X^{-1} \underline{r}_{X,\theta}.$$

Thus we see that, by virtue of its handling of bias, the optimal optimal affine estimator has MSE that will never exceed $\text{var}(\theta)$, the MSE of the constant estimator.

5.3.1 SUPERPOSITION PROPERTY OF LINEAR/AFFINE ESTIMATORS

Let ψ and ϕ be two random variables. Then, as statistical expectation is a linear operator, the best linear (affine) estimator of the sum $\theta = \psi + \phi$ given $\underline{X}$ is

$$\hat{\theta} = \hat{\psi} + \hat{\phi}, \tag{60}$$

where $\hat{\psi}$ and $\hat{\phi}$ are the best linear (affine) estimators of ψ given $\underline{X}$ and of ϕ given $\underline{X}$, respectively.

5.4 GEOMETRIC INTERPRETATION: ORTHOGONALITY CONDITION AND PROJECTION THEOREM

There is a deeper geometrical interpretation of the structure of affine or linear minimum mean squared error estimators. To get at this geometrical interpretation we recast the affine estimation problem into a linear approximation problem in a vector space. For the reader who has only a dim memory of vector spaces we provide a quick review in the Appendix, Sec. 5.10.

5.4.1 LINEAR MINIMUM MSE ESTIMATION REVISITED

The key to embedding this problem into a vector space is to identify the right space for the approximation problem. There are two spaces that we need to keep in mind: the space $\mathcal{H}$ containing quantities we wish to approximate, e.g., θ, and the space $\mathcal{S}$, called the solution subspace, in which we construct the approximation, e.g., linear combinations of the X_i's. The problem then reduces to finding a linear combination of vectors in $\mathcal{S}$ that is closest to the quantity we wish to approximate in $\mathcal{H}$. For the machinery to work it is absolutely required that $\mathcal{S} \subset \mathcal{H}$. Once we identify these spaces it only remains to construct an inner product that induces the proper norm that expresses approximation error as the MSE.

As in the min MSE problem we are attempting to approximate the scalar random variable θ with a linear combination of the measured random variables $X_1, \ldots, X_n$ it makes sense to define $\mathcal{H}$ as the space of all scalar zero mean random variables and $\mathcal{S}$ as the linear span span$\{X_1, \ldots, X_n\}$ of the measurements. For technical reasons we will require that all random variables in $\mathcal{H}$ have finite second moment - otherwise one may end up with vectors with infinite norms and nonsensical approximations of infinity. The MSE between two vectors, i.e., random variables, $\eta, \nu \in \mathcal{H}$ can then be adopted as the squared norm

$$\|\eta - \nu\|^2 = E[(\eta - \nu)^2],$$

which is induced by the inner product

$$\langle \eta, \nu \rangle = E[\eta\nu].$$

Since $\hat{\theta} = \underline{h}^T \underline{X} = \sum_{i=1}^n h_i X_i$ is in $\mathcal{S}$ the linear minimum MSE estimate of θ is the vector $\hat{\theta} \in \mathcal{S}$ which minimizes the norm squared $\|\theta - \hat{\theta}\|^2$. Application of the projection theorem thus asserts the following (see also Fig. 31):

Linear estimator projection theorem: the best linear estimator of θ based on $X_1, \ldots, X_n$ is the projection of θ onto $\mathcal{S} = \text{span}\{X_1, \ldots, X_n\}$.

By the orthogonality condition of the projection theorem, the best linear estimator $\hat{\theta}$ must satisfy

$$\langle \theta - \hat{\theta}, u \rangle = 0, \qquad \text{for all } u \in \mathcal{S}$$

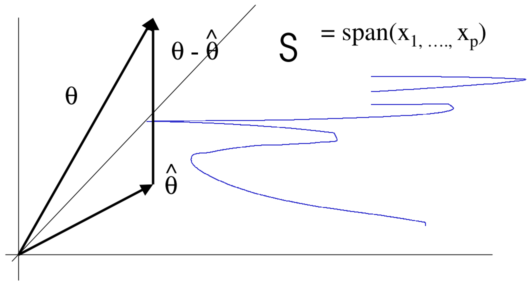


Figure 31: *The orthogonality condition for the best linear estimator $\hat{\theta}$ of a random variable θ given $X_1, \ldots, X_n$*

Equivalently, if $u_1, \ldots, u_{n'}$ is any basis for $\mathcal{S}$:

$$\langle \theta - \hat{\theta}, u_i \rangle = 0, \qquad i = 1, \ldots, n'. \tag{61}$$

When $\{X_i\}_{i=1}^n$ are linearly independent the dimension of $\mathcal{S}$ is equal to n and any basis must have $n' = n$ linearly independent elements.

Now we simply have to adopt a particular basis to find the form of the best linear estimator. Perhaps the most natural basis is the set of measurements themselves $u_i = X_i$, $i = 1, \ldots, n$, (assuming that they are linearly dependent) and, concatenating the $n' = n$ equations (61) into a row vector we obtain

$$E[(\theta - \hat{\theta})\underline{X}^T] = \underline{0}.$$

Equivalently,

$$E[(\theta - \underline{h}^T \underline{X})\underline{X}^T] = \mathbf{M}_{X,\theta}^T - \underline{h}^T \mathbf{M}_X = \underline{0}.$$

As linear independence of the X_i's implies that $\mathbf{M}_X = E[\underline{X}\underline{X}^T]$ is invertible, this yields the identical solution to the optimal coefficients of the linear estimator obtained above: $\underline{h} = \mathbf{M}_X^{-1}\mathbf{M}_{X,\theta}$.

It turns out that a second application of the orthogonality condition yields an immediate expression for minimum MSE:

$$\begin{aligned}
\|\theta - \hat{\theta}\|^2 &= \langle \theta - \hat{\theta}, \theta - \hat{\theta} \rangle \\
&= \langle \theta - \hat{\theta}, \theta \rangle - \langle \theta - \hat{\theta}, \underbrace{\hat{\theta}}_{\in \mathcal{S}} \rangle \\
&= \langle \theta - \hat{\theta}, \theta \rangle \\
&= E[\theta^2] - \underline{h}^T \mathbf{M}_{X,\theta} = E[\theta^2] - \mathbf{M}_{X,\theta}^T \mathbf{M}_X^{-1} \mathbf{M}_{X,\theta},
\end{aligned}$$

where in the second to last line we have used the fact that the optimal error is orthogonal to any vector in $\mathcal{S}$.

5.4.2 AFFINE MINIMUM MSE ESTIMATION

The affine minimum MSE estimation problem is also easily cast into a vector space minimum norm problem. One way to do this is to subtract the mean from the parameter and subtract the mean from the measurements and proceed as in linear estimation - adding the parameter mean back into the solution at the end. A more direct approach is to include the degenerate constant random variable "1" into the measurement vector (we can always add a virtual sensor to the measurement system that measures a nonrandom constant!). To see how this would work first re-express the affine estimator equation as

$$\begin{aligned} \hat{\theta} &= \underline{h}^T \underline{X} + b \\ &= [\underline{h}^T, b] \begin{bmatrix} \underline{X} \\ 1 \end{bmatrix}. \end{aligned}$$

We now identify the solution subspace

$$\mathcal{S} := \text{span}\{X_1, \ldots, X_n, 1\},$$

which gives the following affine projection theorem:

Affine projection theorem: the best affine estimator of θ based r.v.s $X_1, \ldots, X_n$ is the projection of θ onto span$\{X_1, \ldots, X_n, 1\}$.

We leave it to the reader to verify that the aplication of the orthogonality condition to the projection theorem gives the same solution that we derived before.

5.4.3 OPTIMALITY OF AFFINE ESTIMATOR FOR LINEAR GAUSSIAN MODEL

Introduce the addition assumption that $\underline{X}$, θ are jointly Gaussian distributed. Then the minimum MSE estimator $\underline{\hat{\theta}} = \underline{\hat{\theta}}(\underline{X})$ is in fact affine:

$$E[\theta|\underline{X}] = E[\theta] + \underline{r}_{X,\theta}^T \mathbf{R}_X^{-1}(\underline{X} - E[\underline{X}]).$$

One way to show this is to simply compute the conditional mean estimator and verify that it is in fact of the form of the affine estimator above. We take a different approach. Without loss of generality, let's specialize to the case of zero mean θ and $\underline{X}$. Let $\hat{\theta}_l$ be the LMMSE estimator, which is identical to the affine estimator in this case. From the linear projection theorem we know that the optimal estimator error is orthogonal to the measurements

$$E[(\theta - \hat{\theta}_l)\underline{X}] = \underline{0}$$

However, since $\theta - \hat{\theta}_l$ is a linear combination of Gaussian r.v.s it is itself Gaussian. Furthermore, since Gaussian r.v.s that are orthogonal are in fact independent r.v.'s

$$E[(\theta - \hat{\theta}_l)|\underline{X}] = E[(\theta - \hat{\theta}_l)] = 0.$$

Therefore, as $\hat{\theta}_l$ is a function of $\underline{X}$ we have

$$0 = E[(\theta - \hat{\theta}_l)|X] = E[\theta|X] - \hat{\theta}_l,$$

or

$$E[\theta|\underline{X}] = \hat{\theta}_l$$

Which establishes the desired result.

5.5 BEST AFFINE ESTIMATION OF A VECTOR

When the parameter $\underline{\theta} = [\theta_1, \ldots, \theta_p]^T$ is a vector it turns out that our previous results for scalar θ generalize very easily if we adopt the sum of the component MSEs as our error criterion. Define the prior mean vector $E[\underline{\theta}]$ and the cross-correlation matrix

$$\mathbf{R}_{X,\underline{\theta}} = \mathrm{cov}(\underline{X}, \underline{\theta}) = E[(\underline{x} - E[\underline{X}])(\underline{\theta} - E[\underline{\theta}])^T].$$

The sum MSE criterion is defined as

$$\begin{aligned} \mathrm{MSE}(\hat{\underline{\theta}}) &= \sum_{i=1}^{p} \mathrm{MSE}(\hat{\theta}_i) \\ &= \sum_{i=1}^{p} E|\theta_i - \hat{\theta}_i|^2 = \mathrm{trace}\left(E[(\underline{\theta} - \hat{\underline{\theta}})\ (\underline{\theta} - \hat{\underline{\theta}})^T]\right). \end{aligned}$$

Let the affine estimator $\hat{\theta}_i$ of the i-th component of $\underline{\theta}$ be defined by

$$\hat{\theta}_i = \underline{h}_i^T \underline{X} + b_i, \quad i = 1, \ldots, p.$$

Define the affine vector estimator

$$\begin{aligned} \hat{\underline{\theta}} &= [\hat{\theta}_1, \ldots, \hat{\theta}_p]^T = \mathbf{H}^T \underline{X} + \underline{b} \\ \mathbf{H} &= [\underline{h}_1, \ldots, \underline{h}_p]. \end{aligned}$$

The affine minimum MSE vector estimation problem is to find $\mathbf{H}, \underline{b}$ to minimize the sum MSE denoted as $\mathrm{MSE}(\mathbf{H}, \underline{b})$.

The solution is the optimal vector affine estimator

$$\hat{\underline{\theta}} = E[\underline{\theta}] + \mathbf{R}_{\underline{\theta},X} \mathbf{R}_X^{-1} (\underline{X} - E[\underline{X}]). \tag{62}$$

The derivation of this result relies on the fact that each pair $\underline{h}_i$ and b_i appears separately in each of the summands of $\mathrm{MSE}(\hat{\underline{\theta}})$. Hence the minimization of MSE is equivalent to the uncoupled minimization of each $\mathrm{MSE}(\hat{\theta}_i)$.

$$\min_{\hat{\underline{\theta}}} \mathrm{MSE}(\mathbf{H}, \underline{b}) = \sum_{i=1}^{p} \min_{\underline{h}_i, b_i} \mathrm{MSE}(\underline{h}_i, b_i).$$

Therefore the minimum MSE solution is simply the concatenation of the optimal scalar affine estimators of each θ_i:

$$\begin{bmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_p \end{bmatrix} = \begin{bmatrix} E[\theta_1] \\ \vdots \\ E[\theta_p] \end{bmatrix} + \begin{bmatrix} \underline{r}_{\theta_1,X} \mathbf{R}_X^{-1} (\underline{X} - E[\underline{X}]) \\ \vdots \\ \underline{r}_{\theta_p,X} \mathbf{R}_X^{-1} (\underline{X} - E[\underline{X}]) \end{bmatrix},$$

which is equivalent to (62).

We can express the resultant minimum sum MSE as

$$\mathrm{MSE}_{\min} = \mathrm{trace}\left(\mathbf{R}_\theta - \mathbf{R}_{\underline{\theta},X} \mathbf{R}_X^{-1} \mathbf{R}_{X,\underline{\theta}}\right).$$

5.5.1 LINEAR ESTIMATION EXAMPLES

Here we apply the above min MSE results to several examples.

Example 22 *Min MSE Linear Prediction*

In linear prediction one assumes that one measures a segment $\{X_{k-p}, \ldots, X_{k-1}\}$ of a time sequence of measurements $\{X_i\}_{i=\infty}^{\infty}$, also called a time series, and the objective is to form a linear p-th order 1-step predictor of the form

$$\hat{X}_k = \sum_{i=1}^{p} a_i X_{k-i}.$$

We will assume that $\{X_i\}_i$ a zero mean wide sense stationary (w.s.s.) random sequence with autocorrelation function

$$r(k) := E[X_i X_{i-k}]$$

The problem is to find the predictor coefficients $\underline{a} = [a_1, \ldots, a_p]^T$ that minimize the mean squared prediction error: $\mathrm{MSE}(\underline{a}) = E[(X_k - \hat{X}_k)^2]$.

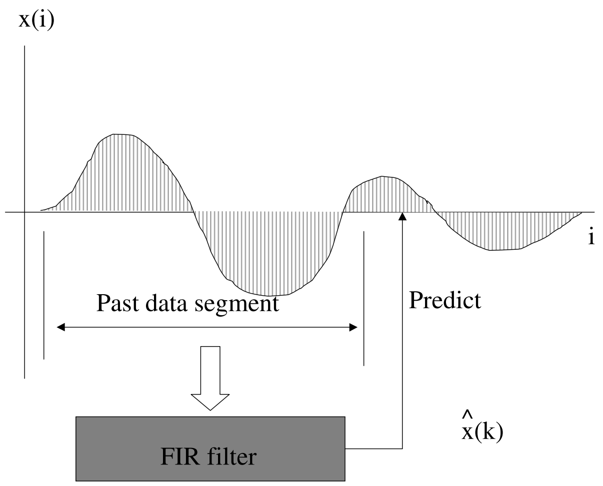


Figure 32: *Linear predictor as a FIR filter.*

To solve for the optimal predictor coefficents identify $\theta = X_k$ as the random scalar parameter, $\underline{X}$ as the time segment measured, and $\underline{h} = \underline{a}$ as the coefficient vector to be determined. We solve the problem in two steps:

Step 1: Rewrite predictor equation in vector form

$$\hat{X}_k = \underline{a}^T \underline{X}$$

where

$$\underline{X} = [X_{k-1}, \ldots, X_{k-p}]^T, \quad \underline{a} = [a_1, \ldots, a_p]^T$$

Step 2: Express orthogonality condition as

$$E[(X_k - \underline{a}^T \underline{X})X_{k-i}] = 0, \quad i = 1, \ldots, p$$

or concatenation into a row vector gives

$$\underline{0}^T = \begin{bmatrix} E[(X_k - \underline{a}^T \underline{X})X_{k-1}] \\ \vdots \\ E[(X_k - \underline{a}^T \underline{X})X_{k-p}] \end{bmatrix} = E[(X_k - \underline{a}^T \underline{X})\underline{X}^T].$$

This specifies the optimal predictor coefficients $\underline{a} = \underline{\hat{a}}$ as

$$\underline{\hat{a}} = \mathbf{R}^{-1}\underline{r}$$

where we have defined the correlation vector:

$$\underline{r}^T = [r_1, \ldots, r_p] = E[\underline{X}X_k],$$

and the (Toeplitz) covariance matrix

$$\mathbf{R} = ((r_{(i-j)}))_{i,j=1,p} = E[\underline{X}\ \underline{X}^T].$$

Finally, the predictor has minimum MSE

$$\begin{aligned} \mathrm{MSE}_{\min} &= \langle X_k - \underline{\hat{a}}^T \underline{X}, X_k \rangle \\ &= r_0 - \underline{\hat{a}}^T \underline{r} \\ &= r_0 - \underline{r}^T \mathbf{R}^{-1} \underline{r} \end{aligned}$$

Relation: The optimal linear predictor can be related to a so-called autoregressive order p (AR(p)) model for $\{X_i\}$.

To see this define residual prediction error $V_k = X_k - \hat{X}_k$. Then, obviously

$$X_k = \hat{X}_k + V_k = \sum_{i=1}^{p} a_i X_{k-i} + V_k$$

When V_k is w.s.s. white noise the above representation is called an AR(p) model for $\{X_i\}$.

Example 23 *An Inverse Problem*

Assume a measurement model:

$$\underline{X} = \mathbf{A}\underline{\theta} + \underline{N}$$

where

* $\underline{X} = [X_1, \ldots, X_m]^T$: random measurements

* $\underline{\theta} = [\theta_1, \ldots, \theta_p]^T$: unknown random parameters

* $\underline{N} = [n_1, \ldots, n_m]^T$: zero mean measurement noise with covariance $\mathbf{R}_N$

* $\underline{\theta}, \underline{N}$ uncorrelated

* $\mathbf{A}$: a known $m \times p$ matrix

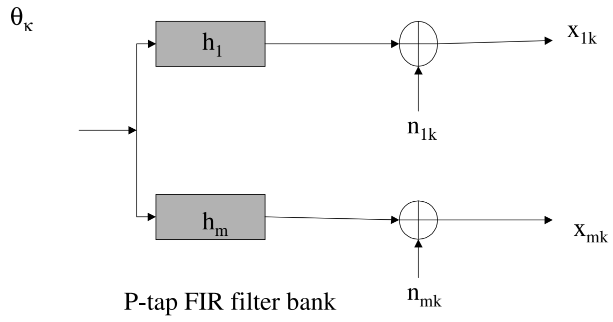


Figure 33: *Block diagram for inverse problem*

The problem is to find an affine min MSE estimator $\hat{\underline{\theta}}$ of $\underline{\theta}$.

Solution: this directly follows from our vector minimum MSE estimation results:

$$\hat{\underline{\theta}} = E[\underline{\theta}] + \mathbf{R}_{\underline{\theta},X}\mathbf{R}_X^{-1}(\underline{X} - E[\underline{X}]).$$

It remains to determine the form of the optimal affine estimator in terms of $\mathbf{A}$ and $\mathbf{R}_N$.

$$\begin{aligned}
E[\underline{X}] &= E[\mathbf{A}\underline{\theta} + \underline{N}] = \mathbf{A}E[\underline{\theta}] \\
\mathbf{R}_X &= \mathrm{cov}(\underbrace{\mathbf{A}\underline{\theta} + \underline{N}}_{\text{uncorrelated}}) = \mathbf{A}\mathbf{R}_\theta\mathbf{A}^T + \mathbf{R}_N \\
\mathbf{R}_{X,\underline{\theta}} &= \mathrm{cov}((\mathbf{A}\underline{\theta} + \underline{N}), \underline{\theta}) = \mathbf{A}\mathbf{R}_{\underline{\theta}}.
\end{aligned}$$

Thus we obtain the final result:

$$\hat{\underline{\theta}} = E[\underline{\theta}] + \mathbf{R}_{\underline{\theta}}\mathbf{A}^T[\mathbf{A}\mathbf{R}_{\underline{\theta}}\mathbf{A}^T + \mathbf{R}_N]^{-1}(\underline{X} - \mathbf{A}E[\underline{\theta}]),$$

and the resultant minimum sum MSE is

$$\mathrm{MSE}_{\min} = \mathrm{trace}\left(\mathbf{R}_\theta - \mathbf{R}_\theta\mathbf{A}^T[\mathbf{A}\mathbf{R}_\theta\mathbf{A}^T + \mathbf{R}_N]^{-1}\mathbf{A}\mathbf{R}_\theta\right)$$

Remarks:

1. When $\mathbf{R}_N$ dominates $\mathbf{A}\mathbf{R}_\theta\mathbf{A}^T$: $\mathrm{MSE}_{\min} \approx \mathrm{trace}\mathbf{R}_{\underline{\theta}}$
2. When $\mathbf{A}\mathbf{R}_{\underline{\theta}}\mathbf{A}^T$ dominates $\mathbf{R}_N$ *and* $\mathbf{A}$ *is full rank*: $\mathrm{MSE}_{\min} \approx 0$.

5.6 NONSTATISTICAL LEAST SQUARES (LINEAR REGRESSION)

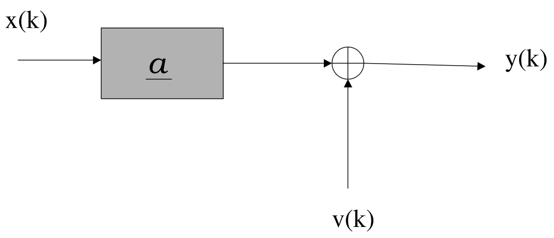
In some cases one does not have a good enough model to compute the ensemble averages, e.g. $\mathbf{R}$ and $\mathbf{R}_{X\theta}$, required for implementation of the linear minimum MSE estimators discussed above.

In these cases one must resort to training data to estimate these ensemble averages. However, a natural question arises: to what extent is it optimal to simply substitute empirical averages into the formulas derived above? The answer depends of course on our definition of optimality. Non-statistical least squares is a new formulation of this problem for which the optimal solutions turn out to be the same form as our previous solutions, but with empirical estimates substituted for $\mathbf{R}$ and $\mathbf{R}_{X,\theta}$.

Assume that a pair of measurements available $(n \geq p)$

$$y_i, \ \ \underline{x}_i = [x_{i1}, \ldots, x_{ip}]^T, \quad i = 1, \ldots, n.$$

x_{ip} could be equal to x_{i-p} here, but this is not necessary.



System diagram for regression model

Figure 34: *System identification block diagram for linear regression*

Postulate an “input-output” relation:

$$y_i = \underline{x}_i^T \underline{a} + v_i, \quad i = 1, \ldots n$$

* y_i is response or output or dependent variable

* $\underline{x}_i$ is treatment or input or independent variable

* $\underline{a}$ is unknown $p \times 1$ coefficient vector to be estimated

$$\underline{a} = [a_1, \ldots, a_p]^T$$

Objective: find linear least squares estimator $\hat{\underline{a}}$ of $\underline{a}$ that minimizes sum of squared errors

$$\mathrm{SSE}(\underline{a}) = \sum_{i=1}^{n} (y_i - \underline{x}_i^T \underline{a})^2$$

Equivalent $n \times 1$ vector measurement model:

$$
\begin{aligned}
\begin{bmatrix} y_1, \\ \vdots \\ y_n \end{bmatrix} &= \begin{bmatrix} \underline{x}_1^T \\ \vdots \\ \underline{x}_n^T \end{bmatrix} \underline{a} + \begin{bmatrix} v_1, \\ \vdots \\ v_n \end{bmatrix} \\
\underline{y} &= \mathbf{X}\underline{a} + \underline{v},
\end{aligned}
$$

where $\mathbf{X}$ is a non-random $n \times p$ input matrix.

The estimation criterion is

$$
\mathrm{SSE}(\underline{a}) = (\underline{y} - \mathbf{X}\underline{a})^T(\underline{y} - \mathbf{X}\underline{a})
$$

Solution to LLSE of $\underline{a}$:

Step 1. Identify vector space containing $\underline{y}$: $\mathcal{H} = \mathbb{R}^n$

Inner product: $\langle \underline{y}, \underline{z} \rangle = \underline{y}^T \ \underline{z}$

Step 2. Identify solution subspace containing $\mathbf{X}\underline{a}$

$$
\mathcal{S} = \mathrm{span}\{\text{columns of } \mathbf{X}\}
$$

which contains vectors of form

$$
\mathbf{X}\underline{a} = \sum_{k=1}^{p} a_k \left[\ x_{1k}, \ldots, x_{nk} \ \right]^T
$$

Step 3. apply projection theorem

Orthogonality Condition: the best linear estimator $\underline{\hat{a}}$ satisfies

$$
\langle \underline{y} - \mathbf{X}\underline{\hat{a}}, \underline{u}_i \rangle = 0, \qquad i = 1, \ldots, n
$$

where $\underline{u}_i$ are columns of $\mathbf{X}$, or equivalently

$$
\begin{aligned}
\underline{0}^T \qquad &= (\underline{y} - \mathbf{X}\underline{\hat{a}})^T \mathbf{X} \\
&= \underline{y}^T \underline{\mathbf{X}} - \underline{\hat{a}}^T \ \mathbf{X}^T \mathbf{X}
\end{aligned}
$$

or, if $\mathbf{X}$ has full column rank p then $\mathbf{X}^T\mathbf{X}$ is invertible and

$$
\begin{aligned}
\underline{\hat{a}} &= [\mathbf{X}^T\mathbf{X}]^{-1} \ \mathbf{X}^T \underline{y} \\
&= [n^{-1}\mathbf{X}^T\mathbf{X}]^{-1} \ [n^{-1}\mathbf{X}^T]\underline{y} \\
&= \hat{\mathbf{R}}_x^{-1} \ \underline{\hat{r}}_{xy}.
\end{aligned}
$$

Here

$$
\hat{\mathbf{R}}_x \stackrel{\mathrm{def}}{=} \frac{1}{n} \sum_{i=1}^{n} \underline{x}_i \ \underline{x}_i^T, \qquad \underline{\hat{r}}_{xy} \stackrel{\mathrm{def}}{=} \frac{1}{n} \sum_{i=1}^{n} \underline{x}_i \ y_i
$$

We next specify the projection operator form of predicted output response

$$\underline{\hat{y}} = \mathbf{X}\underline{\hat{a}}$$

which, using above, can be represented as the orthogonal projection of $\underline{y}$ onto $\mathcal{S}$

$$\begin{aligned}\underline{\hat{y}} &= \mathbf{X}\underline{\hat{a}} \\ &= \mathbf{X}[\mathbf{X}^T\mathbf{X}]^{-1}\,\mathbf{X}^T\underline{y} \\ &= \underbrace{\mathbf{X}[\mathbf{X}^T\mathbf{X}]^{-1}\mathbf{X}^T}_{\text{orthog. projection}}\;\underline{y}\end{aligned}$$

Properties of orthogonal projection operator:

$$\Pi_{\mathbf{X}} = \mathbf{X}[\mathbf{X}^T\mathbf{X}]^{-1}\mathbf{X}^T$$

Property 1. $\Pi_{\mathbf{X}}$ projects vectors onto column space of $\mathbf{X}$

Define decomposition of $\underline{y}$ into components $\underline{y}_{\mathbf{X}}$ in column space of $\mathbf{X}$ and $\underline{y}_{\mathbf{X}}^{\perp}$ orthogonal to column space of $\mathbf{X}$

$$\underline{y} = \underline{y}_{\mathbf{X}} + \underline{y}_{\mathbf{X}}^{\perp}$$

Then for some vector $\underline{\alpha} = [\alpha_1, \ldots, \alpha_p]^T$

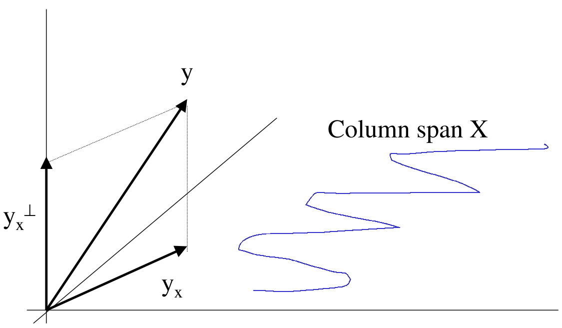


Figure 35: *Column space decomposition of a vector* $\underline{y}$

$$\underline{y}_{\mathbf{X}} = \mathbf{X}\underline{\alpha}, \qquad \mathbf{X}^T\underline{y}_{\mathbf{X}}^{\perp} = \underline{0}$$

We have:

$$\begin{aligned}
\Pi_{\mathbf{X}}\underline{y} &= \Pi_{\mathbf{X}}(\underline{y}_{\mathbf{X}} + \underline{y}^{\perp}) \\
&= \mathbf{X}\underbrace{[\mathbf{X}^T\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{X}}_{=\mathbf{I}}\,\underline{\alpha} + \mathbf{X}[\mathbf{X}^T\mathbf{X}]^{-1}\underbrace{\mathbf{X}^T\underline{y}_{\mathbf{X}}^{\perp}}_{=\underline{0}} \\
&= \mathbf{X}\underline{\alpha} \\
&= \underline{y}_{\mathbf{X}}
\end{aligned}$$

so that $\Pi_{\mathbf{X}}$ extracts the column space component of $\underline{y}$. Thus we can identify $\underline{y}_{\mathbf{X}} = \Pi_{\mathbf{X}}\underline{y}$ so that we have the representation

$$\underline{y} = \Pi_{\mathbf{X}}\underline{y} + \underbrace{(I - \Pi_{\mathbf{X}})\underline{y}}_{\underline{y}_{\mathbf{X}}^{\perp}}$$

It follows immediately that 2. $I - \Pi_{\mathbf{X}}$ projects onto the space orthogonal to span$\{$cols$\mathbf{X}\}$

3. $\Pi_{\mathbf{X}}$ is symmetric and idempotent: $\Pi_{\mathbf{X}}^T\Pi_{\mathbf{X}} = \Pi$

4. $(I - \Pi_{\mathbf{X}})\Pi_{\mathbf{X}} = 0$

Projection operator form of LS estimator gives alternative expression for minimum SSE

$$\begin{aligned}
\text{SSE}_{\min} &= (\underline{y} - \underline{\hat{y}})^T(\underline{y} - \underline{\hat{y}}) \\
&= \underline{y}^T[I - \Pi_{\mathbf{X}}]^T\;[I - \Pi_{\mathbf{X}}]\underline{y} \\
&= \underline{y}^T[I - \Pi_{\mathbf{X}}]\underline{y}
\end{aligned}$$

Example 24 *LS optimality of sample mean*

Measure $\underline{x} = [x_1, \ldots, x_n]^T$

Objective: Find best constant c which minimizes the sum of squares

$$\sum_{k=1}^{n}(x_i - c)^2 = (\underline{x} - c\underline{1})^T(\underline{x} - c\underline{1})$$

where $\underline{1} = [1, \ldots, 1]^T$

Step 1: identify solution subspace

$\mathcal{S}$ is diagonal line: $\{\underline{y} : \underline{y} = a\underline{1},\ a \in \mathbb{R}\}$

Step 2. apply orthogonality condition

$$(\underline{x} - c\underline{1})^T\underline{1} = 0 \iff c = \frac{\underline{x}^T\underline{1}}{\underline{1}^T\underline{1}} = n^{-1}\sum_{k=1}^{n} x_i$$

Example 25 *LLS linear prediction from training sample*

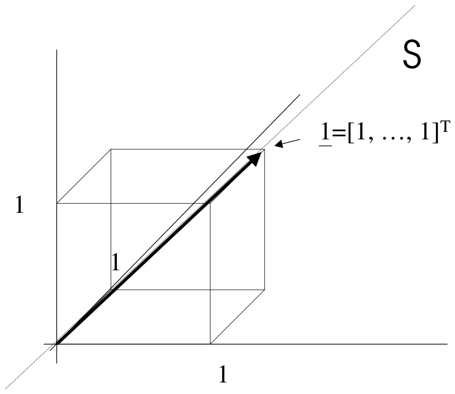


Figure 36: *Diagonal line is solution subspace for LS scalar*

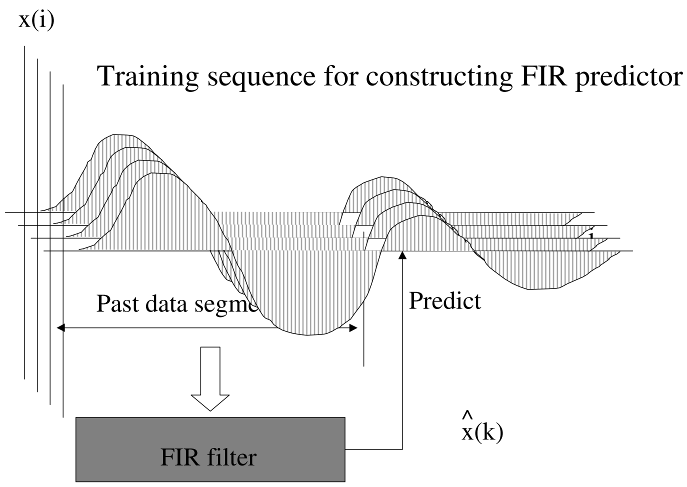


Figure 37: *Construction of LLS predictor from training sequence*

Measurement sequence $\{z_i\}$

Training sequence of $n+p$ samples of z_i

$$\{z_i\}_{i=1}^{p+n}, \quad i=1,\ldots,n$$

Fit an AR(p) model to training sequence

$$z_k = \sum_{i=1}^{p} a_i z_{k-i} + v_k, \quad k=p+1,\ldots,n$$

such that SSE is minimized

$$\mathrm{SSE}(n) = \sum_{k=1}^{n} (z_{k+p} - \sum_{i=1}^{p} a_i z_{k+p-i})^2$$

Solution

Step 1. Identify response variables $y_k = z_k$ and input vectors $\underline{z}_k = [z_{k-1},\ldots,z_{k-p}]^T$.

$$\begin{aligned} \begin{bmatrix} z_{n+p}, \\ \vdots \\ z_{p+1} \end{bmatrix} &= \begin{bmatrix} \underline{z}_{n+p}^T, \\ \vdots \\ \underline{z}_{p+1}^T \end{bmatrix} \underline{a} + \begin{bmatrix} v_{n+p} \\ \vdots \\ v_{p+1} \end{bmatrix} \\ \underline{y} &= \mathbf{X}\underline{a} + \underline{v}, \end{aligned}$$

Step 2. Apply orthogonality condition

The LLS p-th order linear predictor is of the form:

$$\hat{z}_k = \sum_{i=1}^{p} \hat{a}_i z_{k-i}$$

where $\underline{\hat{a}} = [\hat{a}_1,\ldots,\hat{a}_p]^T$ is obtained from formula

$$\underline{\hat{a}} = [\mathbf{X}^T\mathbf{X}]^{-1}\mathbf{X}^T\underline{y} = \hat{\mathbf{R}}^{-1}\,\underline{\hat{r}}$$

and we have defined the sample correlation quantities:

$$\hat{r} = [\hat{r}_1,\ldots,\hat{r}_p]^T$$

$$\hat{\mathbf{R}} = ((\hat{r}(i-j)))_{i,j=1,p}$$

$$\hat{r}_j := n^{-1} \sum_{i=1}^{n} z_{i+p} z_{i+p-j}, \quad j=0,\ldots,p$$

5.7 LINEAR MINIMUM WEIGHTED LEAST SQUARES ESTIMATION

As before assume linear model for input and response variables

$$\begin{aligned}
\begin{bmatrix} y_1, \\ \vdots \\ y_n \end{bmatrix} &= \begin{bmatrix} \underline{x}_1^T, \\ \vdots \\ \underline{x}_n^T \end{bmatrix} \underline{a} + \begin{bmatrix} v_1, \\ \vdots \\ v_n \end{bmatrix} \\
\underline{y} &= \mathbf{X}\underline{a} + \underline{v},
\end{aligned}$$

The linear minimum weighted least squares (LMWLS) estimator $\hat{\underline{a}}$ of $\underline{a}$ minimizes

$$\mathrm{SSE}(\underline{a}) = (\underline{y} - \mathbf{X}\underline{a})^T \; \mathbf{W} \; (\underline{y} - \mathbf{X}\underline{a})$$

where $\mathbf{W}$ is a symmetric positive definite $n \times n$ matrix

Solution to LMWMS problem:

Step 1. Identify vector space containing $\underline{y}$: $\mathcal{H} = \mathbb{R}^n$

Inner product: $\langle \underline{y}, \underline{z} \rangle = \underline{y}^T \; \mathbf{W}\underline{z}$

Step 2. Identify solution subspace $\mathcal{S}$

$$\mathbf{X}\underline{a} = \mathrm{span}\{\text{columns of } \mathbf{X}\}$$

Step 3. apply projection theorem

Orthogonality Condition: the best linear estimator $\hat{\underline{a}}$ satisfies

$$\begin{aligned}
0 &= (\underline{y} - \mathbf{X}\hat{\underline{a}})^T \mathbf{W}\mathbf{X} \\
&= \underline{y}^T \mathbf{W}\underline{\mathbf{X}} - \hat{\underline{a}}^T \; \mathbf{X}^T \mathbf{W}\mathbf{X}
\end{aligned}$$

or, if $\mathbf{X}$ has full column rank p then $\mathbf{X}^T\mathbf{W}\mathbf{X}$ is invertible and

$$\hat{\underline{a}} = [\mathbf{X}^T\mathbf{W}\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{W}\underline{y}$$

5.7.1 PROJECTION OPERATOR FORM OF LMWLS PREDICTOR

The vector $\hat{\underline{y}}$ of least squares predictors $\hat{y}_i = \underline{x}_i^T \hat{\underline{a}}$ of the actual output $\underline{y}$ is

$$\hat{\underline{y}} = \mathbf{X}\hat{\underline{a}}$$

which can be represented as the "oblique" projection of $\underline{y}$ onto $\mathcal{H}$

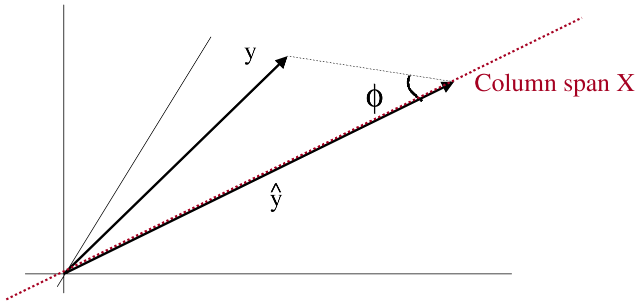


Figure 38: *Oblique projection interpretation of WLS estimator*

$$\underline{\hat{y}} = \underbrace{\mathbf{X}[\mathbf{X}^T\mathbf{W}\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{W}}_{\text{oblique projection } \Pi_{\mathbf{X},\mathbf{W}}} \quad \underline{y}$$

Resultant weighted sum of square error:

$$\begin{aligned} & \mathrm{WSSE}_{\min} \\ = & \ \underline{y}^T \ [\mathbf{I} - \mathbf{X}[\mathbf{X}^T\mathbf{W}\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{W}][\mathbf{I} - \mathbf{X}[\mathbf{X}^T\mathbf{W}\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{W}]^T \ \underline{y} \\ = & \ \underline{y}^T[\mathbf{I} - \Pi_{\mathbf{X},\mathbf{W}}]^T[\mathbf{I} - \Pi_{\mathbf{X},\mathbf{W}}]\underline{y} \end{aligned}$$

ALTERNATIVE INTERPRETATION: LMWLS predictor as linear minimum least squares predictor (unweighted) with preprocessing and postprocessing:

As $\mathbf{W}$ is symmetric positive definite there exists a square root factorization of the form

$$\mathbf{W} = \mathbf{W}^{\frac{1}{2}}\mathbf{W}^{\frac{1}{2}}$$

and

$$\begin{aligned} \underline{\hat{y}} &= \mathbf{W}^{-\frac{1}{2}} \ \underbrace{\mathbf{W}^{\frac{1}{2}}\mathbf{X}[\mathbf{X}^T\mathbf{W}^{\frac{1}{2}} \ \mathbf{W}^{\frac{1}{2}}\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{W}^{\frac{1}{2}}}_{\text{orthog. projector } \Pi_{\mathbf{W}^{\frac{1}{2}}\mathbf{X}}} [\mathbf{W}^{\frac{1}{2}}\underline{y}] \\ &= \mathbf{W}^{-\frac{1}{2}} \ \Pi_{\mathbf{W}^{\frac{1}{2}}\mathbf{X}} \ \mathbf{W}^{\frac{1}{2}}\underline{y} \end{aligned}$$

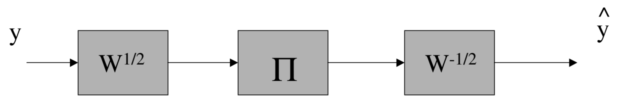


Figure 39: *Interpretation of LMWLS estimator as pre- and postprocessing with orthogonal projection*

Example 26 *Adaptive Linear Prediction*

Now want to fit AR(p) model

$$z_k = \sum_{i=1}^{p} a_i z_{k-i} + v_k, \quad k = 1, 2, \ldots$$

such that at time n we minimize weighted least squares criterion

$$\mathrm{WSSE}(n) = \sum_{k=1}^{n} \rho^{n-k} (z_{k+p} - \sum_{i=1}^{p} a_i z_{k+p-i})^2$$

$\rho \in [0, 1]$ is an exponential forgetting factor
Solution of LMWMS problem:
As before, identify response variables $y_k = z_k$ and input vectors $\underline{x}_k = [z_{k-1}, \ldots, z_{k-p}]^T$.
Also identify weight matrix

$$\mathbf{W} = \begin{bmatrix} \rho^0 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \rho^{n-1} \end{bmatrix}$$

In this way we obtain LMWMS predictor coefficients as

$$\begin{aligned} \underline{\hat{a}} &= [\mathbf{X}^T \mathbf{W} \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{W} \underline{y} \\ &= \hat{\hat{\mathbf{R}}}^{-1} \underline{\hat{\hat{r}}} \end{aligned}$$

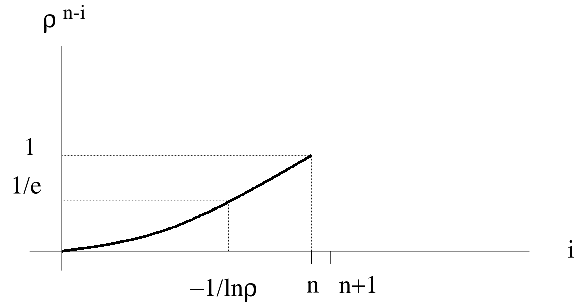


Figure 40: *Exponential forgetting factor applied to past errors for adaptive prediction*

and we have defined the smoothed sample correlation quantities:

$$\hat{\hat{r}} = [\hat{\hat{r}}_1, \ldots, \hat{\hat{r}}_p]^T$$

$$\hat{\hat{\mathbf{R}}} = ((\hat{\hat{r}}(i-j)))_{i,j=1,p}$$

$$\hat{\hat{r}}_j := \sum_{i=1}^{n} \rho^{n-i} z_{i+p} z_{i+p-j}, \quad j = 0, \ldots, p$$

Minimum weighted sum of squared errors (WSSE) is:

$$\text{WSSE}_{\min} = \hat{\hat{r}}_0 - \hat{\hat{r}}^T \hat{\hat{\mathbf{R}}}^{-1} \hat{\hat{r}}$$

5.8 OPTIMALITY OF LMWMS IN THE GAUSSIAN MODEL

Recall that the LMMSE estimator turned out to be globally optimal among arbitrary (linear or non-linear) estimators for a jointly Gaussian measurement and parameter model. Here we show an analogous result for the linear minimum WSSE estimator.

Hypothesize the particular Gaussian model:

$$\underline{Y} = \mathbf{X}\, \underline{a} + \underline{V}$$

where we assume:

* $\underline{V} \sim \mathcal{N}_n(0, \mathbf{R})$

* covariance matrix $\mathbf{R}$ is known

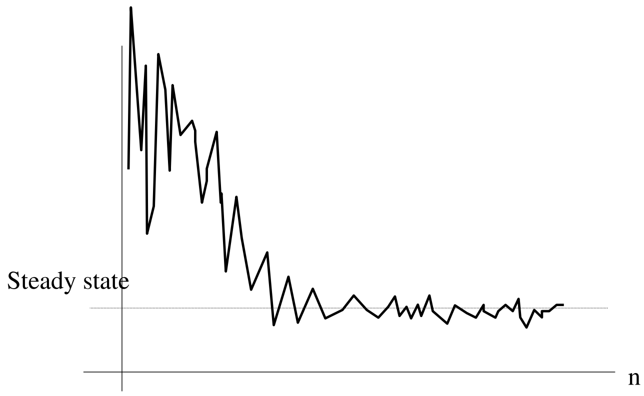


Figure 41: *Typical trajectory of the error criterion for predicting a stationary AR(1) process*

* $\mathbf{X}$ is known non-random matrix of measurements

Under the above hypothesis, for any given $\mathbf{X}$ or $\underline{a}$ the density function of $\underline{Y}$ is multivariate Gaussian

$$f(\underline{y}; \underline{a}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{R}|}} \exp\left(-\frac{1}{2}(\underline{y} - \mathbf{X}\underline{a})^T \mathbf{R}^{-1} (\underline{y} - \mathbf{X}\underline{a})\right).$$

This implies that the maximum likelihood (ML) estimator of $\underline{a}$ is identical to the LMWMS estimator. To see this express

$$\begin{aligned} \underline{\hat{a}}_{ml} &= \text{argmax}_{\underline{a}} \ln f(\underline{Y}; \underline{a}) \\ &= \text{argmin}_{\underline{a}} (\underline{Y} - \mathbf{X}\underline{a})^T \mathbf{R}^{-1} (\underline{Y} - \mathbf{X}\underline{a}). \end{aligned}$$

Hence

$$\hat{Y} = \mathbf{X}\underline{\hat{a}}_{ml} = \mathbf{X}[\mathbf{X}^T\mathbf{R}^{-1}\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{R}^{-1}\underline{Y} = \Pi_{\mathbf{X},\mathbf{W}}\, \underline{Y}.$$

Under the hypothesized model we can evaluate estimator performance by looking to satisfy the condition for equality in the Cramér-Rao bound (CRB)

$$\begin{aligned} (\nabla_{\underline{a}} \ln f)^T &= (\underline{Y} - \mathbf{X}\underline{a})^T\, \mathbf{R}^{-1}\mathbf{X} \\ &= \left(\underbrace{\underline{Y}^T\mathbf{R}^{-1}\mathbf{X}[\mathbf{X}^T\mathbf{R}^{-1}\mathbf{X}]^{-1}}_{\underline{\hat{a}}^T} - \underline{a}^T\right) \underbrace{\mathbf{X}^T\mathbf{R}^{-1}\mathbf{X}}_{K_{\underline{a}}} \end{aligned}$$

We conclude: when $\mathbf{X}$ is nonrandom and noise covariance $\mathbf{R}$ is known to be equal to the LS weighting matrix $\mathbf{W}^{-1}$ then

* the LMWMS estimator $\underline{\hat{a}}$ is unbiased

* the LMWMS estimator is efficient and therefore UMVUE

* recalling property 5 of the CRB in Section 4.5.1, as $\mathbf{K}_{\underline{a}}$ is not a function of $\underline{a}$ the estimator covariance is

$$\mathrm{cov}_{\underline{a}}(\hat{\underline{a}}) = K_{\underline{a}}^{-1} = [\mathbf{X}^T\mathbf{R}^{-1}\mathbf{X}]^{-1} = \hat{\hat{\mathbf{R}}}^{-1}\frac{1}{n}$$

5.9 BACKGROUND REFERENCES

Two classic statistical references for linear estimation are Rao [57] and Anderson [2]. For treatments with more of a signal processing flavor the reader is referred to books by Scharf [60], Van Trees [73] and Kay [36]. The area of control and systems identification have also developed their own distinctive approaches to this problem, see Kailath [32] and Soderstrom and Stoica [66].

5.10 APPENDIX: VECTOR SPACES

For a concise overview of vector spaces in the context of signal processing the reader is referred to Moon and Stirling [49]. For a more advanced treatment with an orientation towards optimization see Luenberger [42].

Definition: $\mathcal{H}$ is a vector space over a scalar field $\mathcal{F}$ if for any elements $x, y, z \in \mathcal{H}$ and scalars $\alpha, \beta \in \mathcal{F}$

1. $\alpha \cdot x + \beta \cdot y \;\in\; \mathcal{H}$ (Closure)
2. $x + (y + z) = (x + y) + z$
3. $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$
4. $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$
5. There is a vector $\phi \in \mathcal{H}$ s.t.: $x + \phi = x$
6. There are scalars $1, 0$ s.t.: $1 \cdot x = x$, $0 \cdot x = \phi$

A normed vector space $\mathcal{H}$ has an inner product $\langle \cdot, \cdot \rangle$ and a norm $\| \cdot \|$ which is defined by $\|x\|^2 = \langle x, x \rangle$ for any $x \in \mathcal{H}$. These quantities satisfy

1. $\langle x, y \rangle = \langle y, x \rangle^*$
2. $\langle \alpha \cdot x, y \rangle = \alpha^* \langle x, y \rangle$
3. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
4. $\|x\| \geq 0$
5. $\|x\| = 0$ iff $x = \phi$
6. $\|x + y\| \leq \|x\| + \|y\|$ (Triangle inequality)
7. $|\langle x, y \rangle| \leq \|x\| \, \|y\|$ (Cauchy-Schwarz inequality)
8. Angle between x, y: $\psi = \cos^{-1}\left(\frac{\langle x,y \rangle}{\|x\| \, \|y\|}\right)$
9. $\langle x, y \rangle = 0$ iff x, y are orthogonal
10. $|\langle x, y \rangle| = \|x\| \, \|y\|$ iff $x = \alpha \cdot y$ for some α

The linear span of vectors $\{x_1, \ldots, x_k\}$ is defined as

$$\mathrm{span}\{x_1, \ldots, x_k\} := \left\{ y : y = \sum_{i=1}^{k} \alpha_i \cdot x_i, \quad \alpha_i \in \mathcal{F} \right\}.$$

A basis for $\mathcal{H}$ is any set of linearly independent vectors $x_1, \ldots x_k$ such that $\text{span}\{x_1, \ldots, x_k\} = \mathcal{H}$

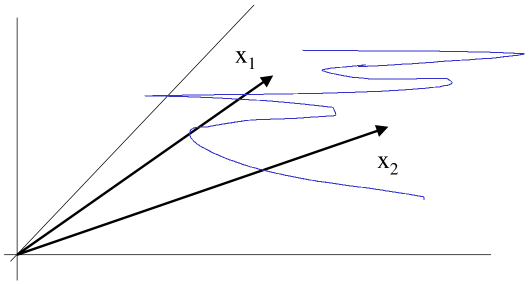


Figure 42: *Illustration of linear span of two vectors in* $\mathbb{R}^3$

The dimension of $\mathcal{H}$ is the number of elements in any basis for $\mathcal{H}$ A linear subspace $\mathcal{S}$ is any subset of $\mathcal{H}$ which is itself a vector space. The projection x of a vector y onto a subspace $\mathcal{S}$ is a vector x that satisfies

$$\langle y - x, u \rangle = 0, \qquad \text{for all } u \in \mathcal{S}$$

The following are some examples of vector spaces:

1. Euclidean p-dimensional space $\mathbb{R}^p$. Identify $\underline{x}$ with x and $\underline{y}$ with y

$$\langle \underline{x}, \underline{y} \rangle = \underline{x}^T \underline{y} = \sum_{i=1}^p x_i y_i$$

A one dimensional subspace: the line

$$\mathcal{S} = \{\underline{y} : \underline{y} = a\underline{v},\ a \in \mathbb{R}\}$$

where $\underline{v} \in \mathbb{R}^p$ is any fixed vector.

2. Complex p-space: $\underline{x} = [x_1, \ldots, x_p]$, $\underline{y} = [y_1, \ldots, y_p]$,

$$\langle \underline{x}, \underline{y} \rangle = \underline{x}^H \underline{y} = \sum_{i=1}^p x_i^* y_i$$

An n-dimensional subspace:

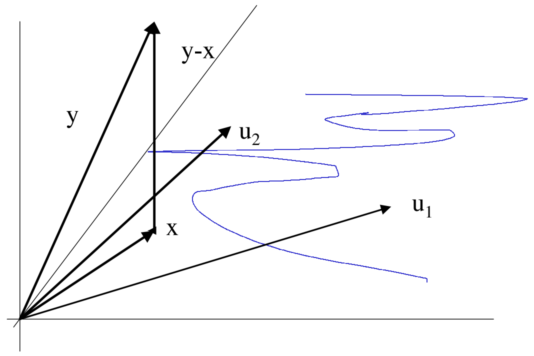


Figure 43: *The projection of a vector x onto a subspace $\mathcal{S}$ in the plane*

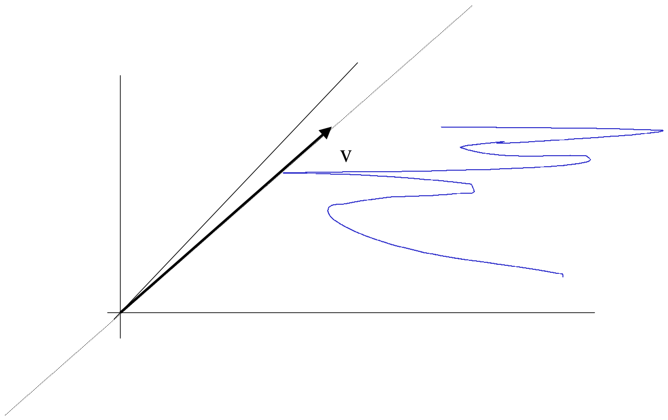


Figure 44: *A line is a one dimensional subspace of* $\mathcal{H} = \mathbb{R}^p$

$$\begin{aligned}\mathcal{S} &= \{\underline{y} : \underline{y} = \sum_{i=1}^{n} a_i \underline{v}_i,\ a_i \in \mathbb{C}\} \\ &= \text{span}\{\underline{v}_1, \ldots, \underline{v}_n\}\end{aligned}$$

where $\underline{v}_i \in \mathbb{C}^p$ are any linearly independent vectors in $\mathcal{H}$.

3. The space of square integrable cts. time functions $x(t)$

$$\langle x, y \rangle = \int x(t) y(t)\, dt$$

A one dimensional subspace: scales of a given function

$$\mathcal{S} = \{g : g(t) = a\, f(t),\ a \in \mathbb{R}\}$$

where $f = f(t)$ is any fixed function in $\mathcal{H}$.

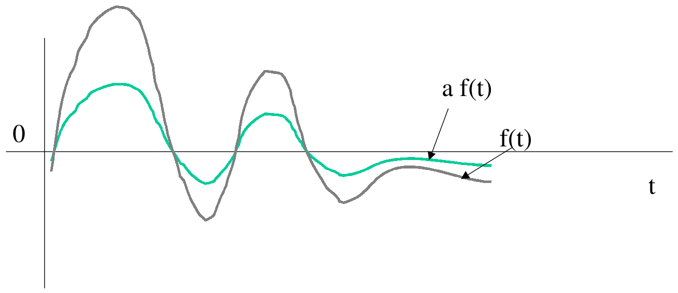


Figure 45: *All scalings of a fixed function is a one dimensional subspace of* $\mathcal{H}$

4. The space of second order real random variables X defined on a sample space. Identify x, y as random variables X, Y:

$$\langle X, Y \rangle = E[XY] = \int_{\Omega} X(\omega) Y(\omega) f(\omega)\, d\omega$$

Ω: sample space of elementary outcomes ω

Q. How to use vector spaces for estimation?

A. Identify $\mathcal{H} = \{Y : Y \text{ a r.v. with } E[|Y|^2] \langle \infty\}$.

Inner product beween two "vectors" in $\mathcal{H}$ is defined as

$$\langle X, Y \rangle := E[XY]$$

(X, Y real r.v.s)

5.11 EXERCISES

5.1 Give a concrete example of two zero mean random variables X and Z for which the linear minimum mean square error estimator of X given Z is equal to 0. Give another concrete example where the overall (linear or non-linear) minimum mean square error estimator is 0.

5.2 Here you explore how linear minimum MSE estimation principles can be applied to non-linear estimation. Given an observed random variable X define the vector of monomials in X:

$$\underline{Y} = [1, X, X^2, \ldots, X^m]^T,$$

where $m \geq 1$ is a positive integer. A non-linear polynomial estimator of order m of another random variable θ given observation X is simply a linear combination of the elements of $\underline{Y}$:

$$\hat{\theta} = \underline{h}^T \underline{Y},$$

where $\underline{h} = [h_0, \ldots, h_m]^T$ is a set of estimation coefficients to be determined. For questions (b) and (c) you may find is useful to use the following facts about a zero mean Gaussian random variable Z: $E[Z^k] = 0$ if k is odd integer while otherwise $E[Z^k] = \sigma_Z^k \cdot (k-1) \cdot (k-3) \cdots 3 \cdot 1$. Also $E[|Z|] = \sigma_Z \sqrt{2/\pi}$.

(a) What is the choice of $\underline{h}$ that minimizes the MSE $E[(\theta - \hat{\theta})^2]$? What is the mean squared error of the resulting optimal non-linear polynomial estimator?

(b) Let $X = sgn(S) + W$, where $sgn(u)$ is the sign of u, S, W are uncorrelated, jointly Gaussian and zero mean. Find the optimal first order ($m = 1$) estimator of S given measurement of X and its resulting minimum MSE. This is the standard optimal linear (affine) estimator we have studied in class.

(c) For the same measurement model for X as in part (b) find the optimal second order ($m = 2$)and third order ($m = 3$) non-linear polynomial estimator of S in terms of σ_w^2 and σ_s^2.

5.3 The least squares method can also be applied to multiple-response linear observation models of the form

$$y_{1k} + \alpha_1 y_{2k} \ldots + \alpha_{p-1} y_{pk} = \beta_1 x_{1k} + \ldots + \beta_q x_{qk} + v_k, \quad k = 1, \ldots, n$$

where $\{y_{1k}, \ldots, y_{pk}\}_k$ are n different observed waveforms (responses) and the α_i. and β_i coefficients are to be determined by minimizing the least squares criterion

$$\text{SSE}(\underline{\alpha}, \underline{\beta}) = \sum_{k=1}^{n} (y_k + \alpha_1 y_{2k} + \ldots + \alpha_{p-1} y_{pk} - \beta_1 x_{1k} - \ldots - \beta_q x_{qk})^2$$

(a) Show that the above observation model is equivalent to the vector model

$$Y[1, \underline{\alpha}^T]^T = X\underline{\beta} + \underline{v}$$

where Y and X are $n \times p$ and $n \times q$ matrices, respectively, $\underline{v}$ is a $n \times 1$ vector of residuals v_k, $\underline{\alpha} = [\alpha_1, \ldots, \alpha_{p-1}]^T$ and $\underline{\beta} = [\beta_1, \ldots, \beta_q]^T$.

(b) Assuming that X has linearly independent columns (full rank) find the least squares estimates $\hat{\underline{\beta}}$ and $\hat{\underline{\alpha}}$ and the resulting minimum SSE. (Hint: first minimize over β then over α)

(c) Assume that the vector $\underline{u} = [1, \underline{\alpha}]^T$ is constrained to have length $\|\underline{u}\| = c$, where $c \geq 1$ is a specified constant. Derive an explicit form for the LS estimators. (Hint: Rayleigh theorem (Ch. 2 or [19]) on minimizing quadratic forms).

(d) The "sensor selection problem" is the following. Fix p' and consider choosing the subset of p' response waveforms $\{y_{i_1,k}\}_{k=1}^n, \ldots, \{y_{i_{p'},k}\}_{k=1}^n$, $i_1, \ldots, i_{p'}$ distinct integers in $1. \ldots, p$, out of the p responses which provide the best fit, i.e. minimize the residuals. Show that the algorithm for solving this sensor selection problem requires solving $p!/(p'!(p-p'))!$ separate least squares problems.

(e) The optimal sensor selection algorithm which you obtained in the previous part of this exercise is of high computational complexity, in the worst case it requires solving approximately $p^{p'}$ least squares problems. Comment on how the solutions to parts (b) or (c) of this exercise could be used to approximate the optimal solution.

5.4 Let the observation have the standard linear model $y_k = \underline{x}_k^T \underline{a} + v_k$, $\quad k = 1, \ldots, n$. We saw in this chapter that when y_k and $\underline{x}_k$ are known and v_k is Gaussian the MLE of $\underline{a}$ is equivalent to the WLSE with weight matrix equal to the covariance matrix of the vector $\underline{v} = [v_1, \ldots, v_n]^T$. In many applications there exist outliers, i.e. a small number of unusually large residual errors v_k, and the Gaussian assumption is not appropriate. Here we treat the case of heavy-tailed distributions of v_k which leads to an estimate of $\underline{a}$ which is more robust to outliers.

(a) Assume that v_k are i.i.d. r.v.s with marginal density $f_v(v)$. Show that the MLE of $\underline{a}$ is

$$\hat{\underline{a}} = \text{argmin}_{\underline{a}} \left\{ \sum_{k=1}^n \log f_v(y_k - \underline{x}_k^T \underline{a}) \right\}$$

(b) Assuming f_v is a smooth function, derive the CR bound on unbiased estimators of $\underline{a}$. Under what conditions is the bound attainable?

(c) Show that for Laplacian noise with $f_v(v) = \frac{\beta}{2} \exp(-\beta|v|)$, $\beta \geq 0$, the MLE reduces to the minimizer of the sum of the absolute errors $|y_k - \underline{x}_k^T \underline{a}|$.

(d) Consider the noise density $f_v(v) = c(\alpha, b) \exp(-v^2/(\alpha^2 + v^2))$, $v \in [-b, b]$, b and α fixed known parameters and c a normalizing constant. Show that the MLE $\hat{\underline{a}}$ can be interpreted as a non-linearly weighted LSE in the sense that it satisfies the "orthogonality condition"

$$\sum_{k=1}^n \lambda_k(\hat{\underline{a}})(y_k - \underline{x}_k^T \hat{\underline{a}}) \underline{x}_k = 0$$

where

$$\lambda_k(\hat{\underline{a}}) = \frac{1}{\alpha^2 + (y_k - \underline{x}_k^T \hat{\underline{a}})^2}$$

(e) The solution to the non-linearly weighted LSE above can be approximated using an "iterative reweighted least squares" technique which consists of approximating the above "orthogonality condition" by implementing the following procedure

i. Initialize $\hat{\underline{a}}_0 = \hat{\underline{a}}$ equal to the standard unweighted LS estimate $\hat{\underline{a}} = [X^T X] X^T \underline{y}$.

ii. Repeat until convergence:

$$\underline{\hat{a}}_{i+1} = [X^T W_i X]^{-1} X^T W_i \underline{y}, i = 1, 2, \ldots$$

where W_i is a diagonal weight matrix with diagonal entries $\lambda_1(\underline{\hat{a}}_i), \ldots, \lambda_n(\underline{\hat{a}}_i)$.

Implement this algorithm in MATLAB and study its convergence for various values of α, b.

5.5 In many applications involving fitting a model to a set of input and output measurements X ($n \times p$) and y ($n \times 1$), not only are the output measurements noisy but the input measurements may also be noisy. In this case the method of Total Least Squares (TLS) [19] is applicable. One formulation of TLS is to model the measurements by

$$y_k = (\underline{x}_k + \underline{\epsilon}_k)^T \underline{a} + v_k, \quad k = 1, \ldots, n$$

where v_k is a zero mean white Gaussian noise with variance σ_v^2 and $\underline{\epsilon}_k$ is an i.i.d. sequence of zero mean Gaussian $p \times 1$ random vectors with diagonal covariance matrix $\sigma_\epsilon^2 \mathbf{I}_p$.

(a) Find the likelihood equation which must be satisfied by the MLE of $\underline{a}$ when σ_v and σ_ϵ are known. To what does your equation reduce when σ_v^2 dominates σ_ϵ^2? What is the ML estimator for $\underline{a}$ in this case?

(b) Show that the MLE of $\underline{a}$ is identical to the standard LS estimator for unknown σ_ϵ.

(c) Find the Fisher information and the CR bound on unbiased estimator covariance for the case of known σ_v and σ_ϵ. Repeat for the case of unknown σ_ϵ. For which of these cases, if any, is the CR bound achievable?

5.6 It is desired to find the linear least sum of squares (LLSS) fit of a complex valued vector $\underline{a}$ to the model

$$y_k = \underline{x}_k^T \underline{a} + v_k, \quad k = 1, \ldots, n$$

where y_k and $\underline{x}_k = [x_{k1}, \ldots, x_{kp}]^T$ are observed. Defining the vector space $\mathcal{H}$ of complex valued n-dimensional vectors with norm $< \underline{y}, \underline{z} >= \underline{y}^H \underline{z}$ ("H" denotes complex conjugate transpose) and vector $\underline{y} = [y_1, \ldots, y_n]^T$ and matrix $X = [\underline{x}_1, \ldots, \underline{x}_n]^T$ (analogously to the case studied in sec. 5.6 of notes). Assume that X is full rank. Using the projection theorem show that the solution to the LLSE problem $\min_{\underline{a}} \|\underline{y} - X\underline{a}\|^2$ is of the form

$$\underline{\hat{a}} = [X^H X]^{-1} X^H \underline{y}$$

with minimum LLSS residual error squared

$$\|\underline{y} - X\underline{\hat{a}}\|^2 = \underline{y}^H [I - X[X^H X]^{-1} X^H] \underline{y}.$$

5.7 This problem applies the solution to the previous exercise. Let the complex observations be given as $X = \{X(0), \ldots, X(N-1)\}$. Hypothesize that $X(k)$ is a damped sinusoid in additive noise:

$$X(k) = a e^{-\alpha k} e^{j 2\pi f_0 k} + Z(k), \quad k \geq 0,$$

where $a \in \mathbb{C}$ is an unknown complex scale factor, $\alpha \geq 0$ is an unknown decay constant, and $f_0 \in [0, \frac{1}{2}]$ is an unknown frequency.

(a) For known α and f_0 show that the least-squares estimator of a which minimizes the sum of the squared residuals $SSE(\underline{a}) = \sum_{k=0}^{N-1} |X(k) - ae^{-\alpha k}e^{j2\pi f_0 k}|^2$ over a has the form (large N):

$$\hat{a} = \mathcal{X}(z_0)\,(1-e^{-2\alpha}),$$

where $\mathcal{X}(z_0) = \sum_{k=0}^{N-1} X(k) z_0^{-k}$ is the Z-transform of X evaluated at the point $z = z_0 = e^{\alpha + j2\pi f_0}$ outside the unit circle. Note that for $\alpha = 0$ this is just the DFT of $X(k)$.

(b) Now for known α but unknown a show that the (non-linear) least-squares estimator for f_0 which minimizes the sum of the squared residuals $SSE(\hat{a}) = \sum_{k=0}^{N-1} |X(k) - \hat{a}e^{-\alpha k}e^{j2\pi f_0 k}|^2$ over f_o is obtained by maximizing the Z-transform of X over the radius e^{α} circle $|z| = e^{\alpha}$:

$$\hat{f}_0 = \text{argmax}_{f_0} |\mathcal{X}(z_0)|^2,$$

and:

$$\hat{a} = \mathcal{X}(e^{\alpha + j2\pi \hat{f}_0})\,(1-e^{-2\alpha}).$$

Note that $\hat{f}_0$ reduces to the location of the highest peak in the magnitude "frequency spectrum" $S(f) = \left|\mathcal{X}(e^{2\pi f})\right|$ of $X(k)$ when α is known to be equal to 0.

(c) Finally for unknown a, α, f_0 show that the non-linear least-squares estimator of α, f_0 is obtained by maximizing the scaled Z-transform of X over the exterior of the unit disk:

$$\hat{f}_0, \hat{\alpha} = \text{argmax}_{f_0, \alpha \geq 0} |\mathcal{X}(z_0)|^2\ (1-e^{-2\alpha}),$$

and:

$$\hat{a} = \mathcal{X}(e^{\hat{\alpha} + j2\pi \hat{f}_0})\,(1-e^{-2\hat{\alpha}}).$$

5.8 It is desired to fit the coefficients α and β to the linear model for the measurements $y_k = \alpha + \beta k + v_k$, $k = 1, \ldots, N$, where v_k is the model error residual to be minimized by suitable chice of α, β. Find the linear least squares estimator for these coefficents (you can leave your solution in the form of a pair of simultaneous equations if you wish).

5.9 It is hypothesized that the relation between a pair of measured variables y_k and x_k is non-linear. A reasonable model for this is

$$y_k = a_0 + a_1 x_k + \ldots + a_p x_k^p + v_k, \quad k = 1, \ldots, n$$

(a) For a single sample ($n = 1$) find the set of coefficients $a_0, \ldots, a_p$ which minimizes the mean squared error $E[(y_k - [a_0 + a_1 x_k + \ldots + a_p x^p])^2]$ under the assumption that y_k and x_k are r.v.'s with known moments $E[y_k x_k^l]$, $l = 0, \ldots, p$, and $E[x_k^l]$, $l = 0, \ldots, 2p$.

(b) Repeat part (a) for the non-statistical least squares estimation error criterion for n samples $\sum_{k=1}^{n} (y_k - [a_0 + a_1 x_k + \ldots + a_p x^p])^2$.

(c) Show that the two estimators found in (a) and (b) become equivalent as $n \to \infty$.

5.10 A sequence of obervations y_k, $k = 1, \ldots, N$ is to be modeled as the sum of two sinusoids

$$y_k = A\cos(\omega_o k) + B\sin(\omega_o k) + v_k$$

where v_k is an error residual, ω_o is known, and A, B are to be determined.

(a) Derive the linear least squares estimators of A, B. Express your result in terms of the real and imaginary parts of the DFT $\mathcal{Y}(\omega) = \sum_{k=1}^{N} y_k e^{-j\omega k}$ of y_k. You may assume that $\sum_{k=1}^{N} \cos(2\omega_o k) = \sum_{k=1}^{N} \sin(2\omega_o k) = 0$.

(b) Now assume that A and B are uncorrelated r.v.s with mean μ_A and μ_B and variance σ^2 and that v_k is zero mean white noise of unit variance uncorrelated with A, B. Derive the affine minimum mean square error estimator of A, B given y_k, $k = 1, \ldots, N$.

(c) Express the result of (b) in terms of the real and imaginary parts of the DFT $\mathcal{Y}(\omega) = \sum_{j=1}^{N} y_k e^{-j\omega k}$ of y_k and compare to the result of part (a)

5.11 In this problem you will explore least squares deconvolution. Available for measurement are the noisy outputs Y_k, $k = 1, \ldots, n$, of a known LTI filter (channel), with known finite impulse response $\{h_k\}_{k=0}^{p}$ and having an unknown input $\{X_k\}$, and measured in additive noise $\{W_k\}$

$$Y_k = \sum_{i=0}^{p} h_i X_{k-i} + W_k, \quad k = 1, \ldots, n$$

The objective is to deconvolve the measurements using the known channel $\{h_k\}$ to recover the input $\{X_k\}$. Assume that $X_k = 0$ for $k \leq 0$.

(a) Show that the above measurement equation can be put in the form

$$\underline{Y} = H\underline{X} + \underline{W},$$

where H is a matrix of impulse responses of the FIR filter. Identify the entries of the vectors $\underline{X}$, $\underline{W}$ and the matrix H.

(b) Assuming that H has linearly independent columns (full rank) find the linear least squares estimate $\hat{\underline{X}}$ which minimizes the sum of squared errors $\sum_{k=1}^{n}(Y_k - h_k * X_k)^2$ (" $*$ " denotes convolution). Give a relation on p, n or $\{h_i\}$ to ensure that H has full rank.

(c) In some cases estimation errors in the recent past are more important than errors in the more distant past. Comment on how you would incorporate this into a weighted linear least squares criterion and find the criterion-minimizing linear estimator $\hat{\underline{X}}$.

(d) A simple model for imprecise knowledge of the channel is

$$\underline{Y} = (H + zI)\underline{X} + \underline{W}$$

where z is a zero mean Gaussian random variable with variance σ^2. Assuming that $\underline{W}$ is zero mean Gaussian random vector with identity covariance (I) find the likelihood function for $\underline{\theta} \stackrel{\text{def}}{=} \underline{X}$ based on the observation $\underline{Y}$. Show that the ML estimator reduces to the linear least squares estimate of part (b) when $\sigma^2 \to 0$.

End of chapter

6 OPTIMAL LINEAR FILTERING AND PREDICTION

In the last chapter linear and affine estimation was explored for the case of a small number n of measurements, or at least small enough so that the $n \times n$ covariance matrix $\mathbf{R}_X$ could be inverted permitting implementation of the linear minimum mean squared error estimator. In this chapter we turn to the case where an online linear predictor is desired, the number n of measurements is increasing over time, and matrix inversion quickly becomes impractical. This situation arises in many signal processing applications where "streaming data" is collected. The approach that is taken for this problem is to model the measurements as coming from a random process with known autocorrelation function (acf). Under some additional assumptions on the acf the linear predictors can be implemented recursively by applying a filter to the data stream. In Wiener filtering, designed for wide sense stationary (wss) processes, this filter is causal and linear time invariant (LTI) while in Kalman-Bucy filtering, designed for non-stationary processes, the filter is linear time varying (LTV).

We will cover the following topics.

* Wiener-Hopf Equations of min MSE filtering for w.s.s. processes

* Non-causal filtering, estimation, and prediction

* Causal LTI prewhitening: spectral factorization

* Causal LTI prediction: the Wiener filter

* Causal LTV prewhitening: the innnovations filter

* Causal LTV prediction: the Kalman-Bucy filter

A word about notation is in order here. Up to now in this text we have used upper case letters for random variables reserving lower case for their realizations. However, it is customary to drop the upper case for random processes and we will do so in this chapter putting the reader at small risk of confusion between realizations, i.e. waveforms, and random processes. Fortunately, second order statistical treatments like that covered here incur fewer accidents due to this kind of abuse of notation.

6.1 WIENER-HOPF EQUATIONS OF OPTIMAL FILTERING

The Wiener filter is useful for linear prediction when one assumes that the underlying random processes are wide sense stationary. The derivation of the filter below requires some facility with z-domain power spectral densities (PSD) for which the reader is referred to the Appendix, Sec. 6.13.

Two zero mean w.s.s. discrete time random processes x and g are of interest to us:

* $x = \{x_k : k \in \mathcal{I}\}$: observed over an index set $\mathcal{I}$

* $g = \{g_k : -\infty < k < \infty\}$: unobserved and to be estimated from x

Objective: estimate a time sample of g, e.g. g_i, by a linear function of the waveform x. Note that g_k plays the role of a random parameter, which we denoted θ_k in previous chapters.

$$\hat{g}_k = \sum_{j \in \mathcal{I}} h(k,j) x_j$$

such that we achieve minimum MSE over choice of linear coefficients $h = \{h(k,j)\}_{k,j}$:

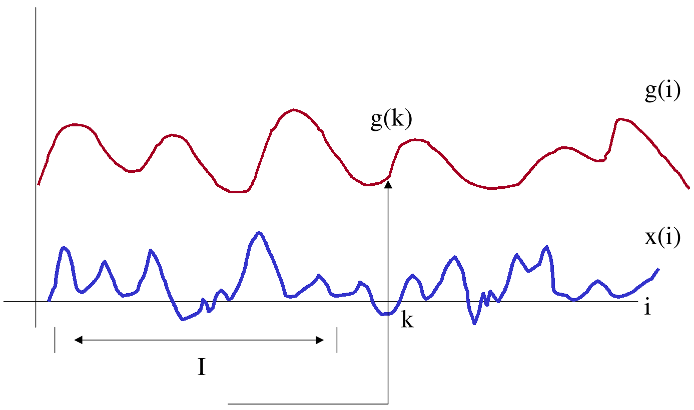


Figure 46: *Two random processes over time. Objective is to predict g_k as a linear function of observation x_i over a time interval $\mathcal{I}$.*

$$\mathrm{MSE}(h) = E[|g_k - \hat{g}_k|^2]$$

Solution: by orthogonality condition

$$E[(g_k - \hat{g}_k)u_i^*] = 0, \qquad i \in I$$

for any basis set $\{u_i\}$ spanning span$\{x_i : i \in \mathcal{I}\}$. The basis $u_i = x_i$, $i \in \mathcal{I}$, will suffice here.

Obtain Wiener-Hopf (WH) equation for optimal filter $h(k,j)$ from orthogonality. condition

$$\begin{aligned}
0 &= E[(g_k - \hat{g}_k)x_i^*] \\
&= E[g_k x_i^*] - \sum_{j \in \mathcal{I}} h(k,j) E[x_j x_i^*] \\
&= r_{gx}(k-i) - \sum_{j \in \mathcal{I}} h(k,j) r_x(j-i), \qquad i \in \mathcal{I}
\end{aligned}$$

When $\mathcal{I}$ is finite this is equivalent to a matrix equation which can be solved as in previous section.

Two cases of interest here:

Case I: $\mathcal{I} = \{-\infty, \ldots, \infty\}$: non-causal estimation (smoothing)

Case II: $\mathcal{I} = \{-\infty, \ldots, k\}$: Causal estimation (filtering)

6.2 NON-CAUSAL ESTIMATION

For $\mathcal{I} = \{-\infty, \ldots, \infty\}$ WH equation becomes

$$r_{gx}(k-i) - \sum_{j=-\infty}^{\infty} h(k,j) r_x(j-i) = 0, \quad -\infty < i < \infty$$

Fact: solution $h(k,j)$ is an LTI filter $h(k-j)$ (Exercise in this chapter).

Take double-sided Z-transform to obtain:

$$\mathcal{P}_{gx}(z) - H(z)\mathcal{P}_x(z) = 0$$

or optimum filter has frequency domain transfer function

$$H(e^{j\omega}) = \frac{\mathcal{P}_{gx}(e^{j\omega})}{\mathcal{P}_x(e^{j\omega})}$$

By invoking the orthogonality condition the minimum MSE can be expressed as follows

$$\begin{aligned} \text{MSE}_{\min} &= E[(g_k - h_k * x_k) g_k^*] \\ &= r_g(0) - h_k * r_{xg}(k)|_{k=0} \end{aligned}$$

Or in frequency domain (recall notation $\mathcal{P}(\omega)$ for $\mathcal{P}(e^{j\omega})$)

$$\begin{aligned} \text{MSE}_{\min} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} [\mathcal{P}_g(\omega) - H(\omega)\mathcal{P}_{xg}(\omega)]\, d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\mathcal{P}_g(\omega) - \frac{\mathcal{P}_{gx}(\omega)\mathcal{P}_{xg}(\omega)}{\mathcal{P}_x(\omega)}\right] d\omega \end{aligned} \tag{63}$$

Example 27 *Wide sense stationary signal in additive noise*

Measurement model

$$x_i = s_i + w_i, \quad -\infty < i < \infty$$

* s, w are uncorrelated w.s.s. random processes
* $g_k = s_k$ to be estimated

Obtain min MSE filter for estimation of s_k and its min MSE

$$H(\omega) = \frac{\mathcal{P}_s(\omega)}{\mathcal{P}_s(\omega) + \mathcal{P}_w(\omega)} = \begin{cases} 1, & \mathcal{P}_s(\omega)/\mathcal{P}_w(\omega) \gg 1 \\ 0, & \mathcal{P}_s(\omega)/\mathcal{P}_w(\omega) \ll 1 \end{cases}$$

$$\text{MSE}_{\min} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\mathcal{P}_s(\omega)\mathcal{P}_w(\omega)}{\mathcal{P}_s(\omega) + \mathcal{P}_w(\omega)} d\omega = \begin{cases} 0, & \min \omega \mathcal{P}_s(\omega)/\mathcal{P}_w(\omega) \gg 1 \\ \text{var}(s_k), & \max_\omega \mathcal{P}_s(\omega)/\mathcal{P}_w(\omega) \ll 1 \end{cases}$$

Problem: $H(\omega)$ is non-negative real so that h_k is symmetric impulse response. This implies that $\hat{s}_k = \sum_{j=-\infty}^{\infty} h(j) x_{k-j}$ depends on future measurements (non-causal)!

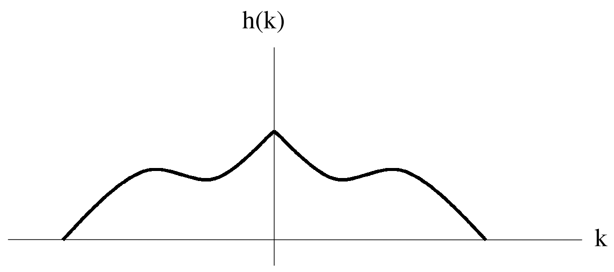


Figure 47: *Min MSE filter for estimating w.s.s. signal in noise has symmetric impulse response.*

6.3 CAUSAL ESTIMATION

Objective: find linear min MSE estimate of g_k based only on past measurements

$$\hat{g}_k = \sum_{j=-\infty}^{k} h(k-j)x_j$$

where h satsifies Wiener-Hopf equations

$$0 = r_{gx}(k-i) - \sum_{j=-\infty}^{k} h(k-j)r_x(j-i), \quad -\infty < i \leq k \tag{64}$$

Let's explicitly constrain filter h_j to be causal: $h(j) = 0$, $j < 0$. Then we have after change of variable (see homework exercises)

$$0 = r_{gx}(l) - \sum_{j=-\infty}^{\infty} h(l-j)r_x(j), \quad l \geq 0 \tag{65}$$

Difficulty: cannot just take z-transform as we did before due to restriction on l. Indeed the equation (65) *does not specify* the value of the difference on the LHS for negative values of l; these values can be *arbitrary* as long as RHS = 0 for $l \geq 0$.

6.3.1 SPECIAL CASE OF WHITE NOISE MEASUREMENTS

One case where solution to WH is simple: $x_i = w_i$ = white noise of unit variance:

$$r_w(k) = \delta_k = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

Wiener-Hopf Eqn becomes

$$r_{gw}(l) - \sum_{j=-\infty}^{\infty} h(l-j) r_w(j) = r_{gw}(l) - h(l), \quad l \geq 0$$

Hence we can specify optimal causal filter as

$$h(k) = \begin{cases} r_{gw}(k), & k \geq 0 \\ 0, & o.w. \end{cases}$$

Or in z-transform domain:

$$H(z) = \{\mathcal{P}_{gw}(z)\}_+ \tag{66}$$

where we have defined truncated z-transform of a time function $b(k)$ with z-transform $B(z)$

$$\{B(z)\}_+ \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} b(k) z^{-k}$$

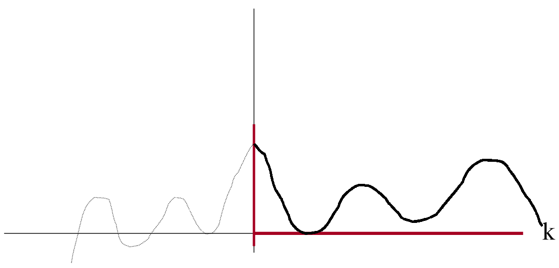


Figure 48: *Truncated z-transform of a function b_k.*

6.3.2 GENERAL CASE OF NON-WHITE MEASUREMENTS

Our derivation of the Wiener filter is based on the approach of Bode and Shannon [8]. The main idea behind this derivation is the following. If we could "prewhiten" x with a filter h_w then we could follow with optimal filter of form (66). This suggests a "prewhitening approach" to solving general problem. However, not just any old whitening filter will do. In keeping with the original objective of causal linear minimum mean square error estimation, the prewhitening filter must itself be causal *and*, to ensure that we lose no information about $\{x_i\}_{i=-\infty}^k$, it must be causally invertible, i.e. we must be able to recover past values of the input $\{x_i\}_{i=-\infty}^k$ from past values of the output.

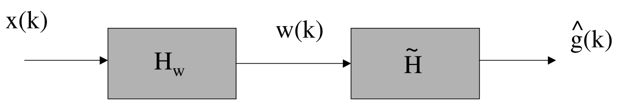


Figure 49: *Solution of the causal estimation problem by a cascade of a prewhitening filter h_w and an optimal wiener filter $\tilde{h}$ for white noise.*

Thus the optimal filter $\tilde{H}$ for whitened measurements is specified by

$$\begin{aligned}\tilde{H}(z) &= \{\mathcal{P}_{gw}(z)\}_+ \\ &= \left\{\mathcal{P}_{gx}(z) H_w(z^{-1})\right\}_+\end{aligned}$$

The filter H_w must satisfy conditions

1. h_w whitens the input process

$$\mathcal{P}_w(z) = H_w(z)H_w(z^{-1})\mathcal{P}_x(z) = 1$$

2. h_w is causal

3. h_w is causally invertible

Definition: A filter (discrete time) h_w with transfer function $H_w(z)$ is causal and causally invertible iff $H_w(z)$ and $1/H_w(z)$ have no singularities outside the unit circle, i.e.,

$$|H_w(z)| < \infty, \quad |z| > 1$$

and similarly,

$$1/|H_w(z)| < \infty, \quad |z| > 1$$

.

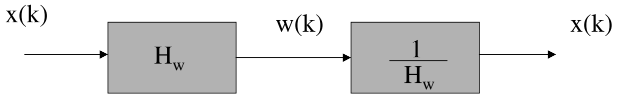


Figure 50: *The cascade of causal filters H_w and $1/H_w$ is causal and all pass thus implying that $\{w_i\}_{i=-\infty}^{k}$ contains same information as $\{x_i\}_{i=-\infty}^{k}$, i.e. w_i is white sufficient statistic.*

6.4 CAUSAL PREWHITENING VIA SPECTRAL FACTORIZATION

Assume x_k is w.s.s. with

* $r_x(k)$ positive definite and summable $\left(\sum_{k=-\infty}^{\infty} |r_x(k)| < \infty\right)$

* rational PSD

$$\mathcal{P}_x(z) = \sum_k r_x(k) z^{-k} = \frac{b(z)}{a(z)}$$

where

$r_x(k) = E[x_i x_{i-k}]$ is acf of x

$b(z) = b_q z^q + \cdots + b_1 z + b_0$

$a(z) = a_p z^p + \cdots + a_1 z + a_0$

For h_w to satisfy whitening condition require

$$H_w(z) H_w(z^{-1}) = \frac{1}{\mathcal{P}_x(z)} \tag{67}$$

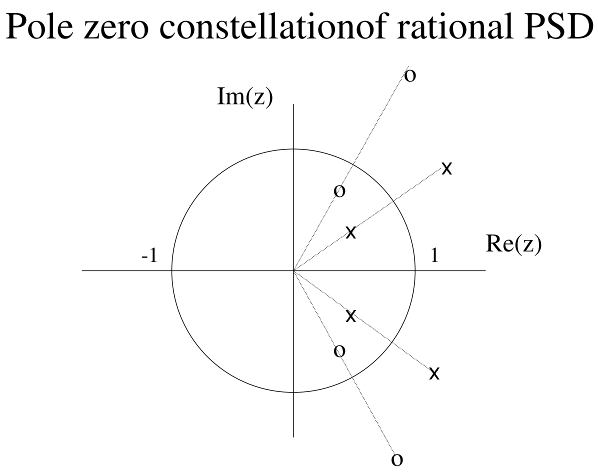


Figure 51: *Pole zero constellation of a rational PSD.*

Next we deal with causality conditions

Any rational PSD can be factored into the form of a ratio of factors of simple first order polynomials

$$\mathcal{P}_x(z) = c\frac{\Pi_{i=1}^q (1 - z/z_{oi})(1 - zz_{oi})}{\Pi_{i=1}^p (1 - z/z_{pi})(1 - zz_{pi})},$$

where c is a real constant and z_{oi}, z_{pi} are zeros and poles of $\mathcal{P}_x(z)$.

This factorization implies the following important properties of rational PSDs

$$\begin{array}{lll} \mathcal{P}_x(z^{-1}) = & \mathcal{P}_x(z), & \text{(symmetric } r_k) \\ \mathcal{P}_x(z^*) = & \mathcal{P}_x^*(z), & \text{(real } r_k) \\ \mathcal{P}_x(z) \rightarrow & \text{no poles on unit circle} & \text{(bounded } r_k) \\ \mathcal{P}_x(z) \rightarrow & \text{zeros on unit circle occur in pairs} & \text{(p.d. } r_k) \end{array}$$

These conditions imply that there exist positive square root factors $\mathcal{P}_x^+(z)$ and $\mathcal{P}_x^-(z)$ which satisfy:

$$\begin{aligned} \mathcal{P}_x(z) &= \mathcal{P}_x^+(z)\mathcal{P}_x^-(z) \\ \mathcal{P}_x^+(z^{-1}) &= \mathcal{P}_x^-(z) \end{aligned}$$

and

$\mathcal{P}_x^+(z)$ has all poles and zeros inside unit circle

$\mathcal{P}_x^-(z)$ has all poles and zeros outside unit circle

Therefore, conclude that the assignment

$$H_w(z) = 1/\mathcal{P}_x^+(z)$$

satisfies whitening condition (67) and $H(z)$, $1/H(z)$ have all their poles inside unit circle

Can identify $1/H_w(z)$ causal synthesis filter for measurement process x_k

$$x_k = h_w^{-1}(k) * w_k$$

where $h_w^{-1}(k)$ is the inverse Z-transform of $1/H_w(z) = \mathcal{P}_x^+(z)$.

This can be useful for simulation of x_k with arbitrary rational PSD from pseudo-random white noise samples w_k.

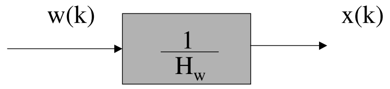


Figure 52: *A representation of x_k with PSD $\mathcal{P}_x$ as output of LTI causal filter $H_w(z) = \mathcal{P}_x^+$ driven by white noise*

6.5 CAUSAL WIENER FILTERING

Putting H_w and $\tilde{H}$ together we obtain formula for the causal Weiner filter

$$\begin{aligned} H(z) &= H_w(z)\,\tilde{H}(z) = \frac{1}{\mathcal{P}_x^+(z)}\left\{\mathcal{P}_{gx}(z)H_w(z^{-1})\right\}_+ \\ &= \frac{1}{\mathcal{P}_x^+(z)}\left\{\frac{\mathcal{P}_{gx}(z)}{\mathcal{P}_x^-(z)}\right\}_+. \end{aligned} \tag{68}$$

A time-domain expression for the minimum mean squared error $\mathrm{MSE}_{\min} = E[(g_k - \hat{g}_k)^2]$ can be simply derived using the orthogonality condition

$$\mathrm{MSE}_{\min} = r_g(0) - \sum_{k=0}^{\infty} h(k) r_{xg}(-k), \tag{69}$$

where $h(k)$ is the inverse Z-transform of $H(z)$ in (68). Unlike the case of non-causal estimation (recall expression (63) there is no simple frequency-domain representation for $\text{MSE}_{\min}$.

Example 28 *Causal prewhitening an AR(1) noise process*

$$x(k) = -ax(k-1) + u(k), \quad (-1 < a < 1)$$

where $u(k)$ is white noise with variance 1.

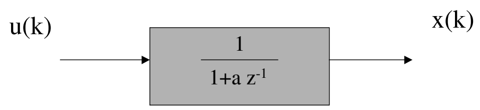


Figure 53: *Synthesis filter of AR(1) process is a single pole IIR.*

First find PSD.

$$\mathcal{P}_x(z) = \underbrace{\frac{1}{(1+az^{-1})}}_{\mathcal{P}_x^+(z)} \underbrace{\frac{1}{(1+az)}}_{\mathcal{P}_x^-(z)}, \qquad r_x(k) = \frac{a^k}{1-a^2}$$

The causal prewhitening filter is FIR

$$H_w(z) = 1 + az^{-1} \Longleftrightarrow h_w(k) = \delta(k) + a\delta(k-1)$$

Can be implemented even without access to infinite past

Example (ctd.) *Prediction of an AR(1) from noiseless observations*

Now we let $g_k = x_{k+\alpha}$ where α is a positive integer. When $\alpha > 0$ this is a prediction problem. In light of the fact that we have just found the prewhitening filter, it remains to find the quantity

$$\left\{ \frac{\mathcal{P}_{gx}(z)}{\mathcal{P}_x^-(z)} \right\}_+$$

to specify the Wiener filter-predictor (68).

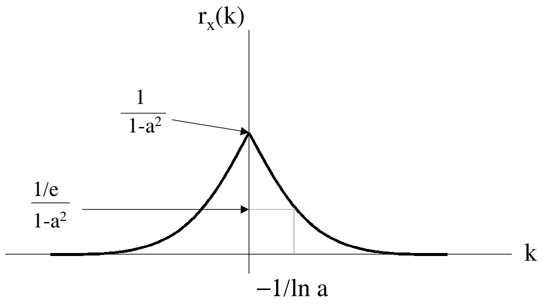


Figure 54: *Auto-correlation function of AR(1) process with AR coefficient a is slowly decaying double sided exponential for $-1 \ll a \leq 0$. (figure k-axis label should be $-1/\ln|a|$).*

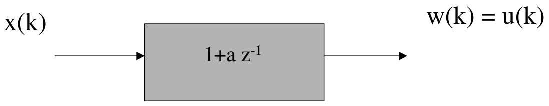


Figure 55: *Causal prewhitening filter for AR(1) process with AR coeffient a is a single tap FIR filter.*

As $r_{gx}(k) = E[x_{i+\alpha}x_{i-k}] = r_x(k+\alpha)$: $\mathcal{P}_{gx}(z) = z^\alpha \mathcal{P}_x(z)$. Hence

$$\begin{aligned}\left\{\frac{\mathcal{P}_{gx}(z)}{\mathcal{P}_x^-(z)}\right\}_+ &= \left\{z^\alpha \mathcal{P}_x^+(z)\right\}_+ \\ &= \left\{z^\alpha/(1+az^{-1})\right\}_+\end{aligned}$$

Now, using the identity (95) derived in the exercises, we see that

$$\left\{z^\alpha/(1+az^{-1})\right\}_+ = (-a)^\alpha/(1+az^{-1})$$

Hence, the Wiener filter-predictor is simply

$$H(z) = \frac{1}{\mathcal{P}_x^+(z)}\left(\frac{(-a)^\alpha}{1+az^{-1}}\right) = (-a)^\alpha$$

which in the time domain gives the optimal predictor as $\hat{x}_{k+\alpha} = (-a)^\alpha x_k$. This just corresponds to scaling the most recent observation and is consistent with x_k being a 1st order Markov sequence so that, for predicting the future, past information is not useful given present information.

Example 29 *Causally prewhitening an AR(1) plus white noise process*

$$x(k) = v_{\mathrm{A}R(1)}(k) + V(k)$$

where

* $v_{\mathrm{A}R(1)}(k)$ is AR(1) with $a = -0.8$

* $V(k)$ is white noise of variance 1/0.36

* $v_{\mathrm{A}R(1)}(k)$ and $V(k)$ are uncorrelated

Using result (94) derived in the Exercises, find PSD as a rational function with double pole and double zero

$$\begin{aligned}\mathcal{P}_x(z) &= \frac{1}{(1-0.8z^{-1})}\frac{1}{(1-0.8z)} + 1/0.36 \\ &= \underbrace{\frac{d(1+bz^{-1})}{(1+az^{-1})}}_{\mathcal{P}_x^+(z)}\ \underbrace{\frac{d(1+bz)}{(1+az)}}_{\mathcal{P}_x^-(z)}\end{aligned}$$

where $a = -0.8, b = -0.5$ and $d = 1/\sqrt{0.225}$.

Unlike previous example, the causal prewhitening filter $h_w(k)$ is now IIR

$$H_w(z) = 1/d\frac{(1+az^{-1})}{(1+bz^{-1})}$$

and thus prewhitening cannot be implemented without access to infinite past.

Note that the synthesis filter $1/H_w(z)$ can be applied to white noise w_k to obtain recursion for x_k with both an autoregressive (AR) component (LHS) and a moving average (MA) component (RHS):

$$x_k + a x_{k-1} = b_1 w_k + b_2 w_{k-1}$$

where $b_1 = d$ and $b_2 = db$. Random processes x_k that satisfy the recursion above are "ARMA(1,1)" process.

Example (ctd.) *Prediction of AR(1) from noisy observations*

Similarly to the previous example we let $\hat{g}(k) = v_{\mathrm{AR}(1)}(k+\alpha)$ where α is a non-negative integer. As the measurement noise $u(k)$ and the AR(1) process $v_{\mathrm{AR}(1)}(k)$ are uncorrelated $\mathcal{P}_{gx}(z) = \mathcal{P}_{gv}(z) = z^\alpha \mathcal{P}_v(z)$ where $\mathcal{P}_v(z)$ is the PSD of $v_{\mathrm{AR}(1)}$ and $\mathcal{P}_{gv}(z)$ is the cross spectral density of g and $v_{\mathrm{AR}(1)}$. Therefore, after substitution of the expression for $\mathcal{P}_x^-$ obtained above,

$$\left\{\frac{\mathcal{P}_{gx}(z)}{\mathcal{P}_x^-(z)}\right\}_+ = \frac{1}{d}\left\{z^\alpha \frac{1}{1+bz}\,\frac{1}{1+az^{-1}}\right\}_+ . \tag{70}$$

Before proceeding further, we will need to express the product of two ratios in $\{\cdot\}_+$ as a sum of two ratios in order to apply the identities (95) and (96). To do this, observe

$$\begin{aligned}\frac{1}{1+bz}\,\frac{1}{1+az^{-1}} &= \frac{z^{-1}}{b+z^{-1}}\,\frac{1}{1+az^{-1}} \\ &= \frac{A}{b+z^{-1}} + \frac{B}{1+az^{-1}}\end{aligned}$$

where A and B are to be determined. Comparing the LHS of the top line to the bottom line of this last equation it is obvious that

$$\begin{aligned}A &= \lim_{z^{-1}\to -b} \frac{z^{-1}}{1+az^{-1}} = -b/(1-ab) \\ B &= \lim_{z^{-1}\to -1/a} \frac{z^{-1}}{b+z^{-1}} = 1/(1-ab)\end{aligned}$$

Thus we have from (70)

$$\begin{aligned}\left\{\frac{\mathcal{P}_{gx}(z)}{\mathcal{P}_x^-(z)}\right\}_+ &= \frac{1}{d(1-ab)}\left\{\frac{z^\alpha}{1+az^{-1}} - \frac{z^{\alpha+1}b}{1+bz}\right\}_+ \\ &= \frac{1}{d(1-ab)}\,\frac{(-a)^\alpha}{1+az^{-1}}\end{aligned}$$

where we have used the identity (96) which shows that only the first additive term in $\{\cdot\}_+$ survives (the second term corresponds to an anticausal component).

Hence, using (68) the Wiener filter-predictor is simply

$$H(z) = \frac{q}{1+bz^{-1}}$$

where $q = \frac{(-a)^\alpha}{d^2(1-ab)}$, which can be implemented in the time domain as the single pole IIR filter recursion

$$\hat{g}(k) = -b\hat{g}(k-1) + q x_k.$$

with $\hat{g}(k) = \hat{v}_{AR(1)}(k+\alpha)$. It can be readily verified that in the limit as the measurement noise $\text{var}(u(k))$ goes to zero, $b \to 0$, $d^2 \to 1$, and $q \to (-a)^\alpha$ so that this IIR predictor filter reduces to the simple Wiener predictor filter of the previous example having no measurement noise.

Derivation of the MSE of the Wiener filter is left as an exercise for the reader.

6.6 CAUSAL FINITE MEMORY TIME VARYING ESTIMATION

The Wiener filter is limited to the cases where the processes g_k and x_k are jointly w.s.s. and the estimating filter is LTI, i.e. access to infinite past is available. In practice, however, this is not the case and we will need to handle the situation for which

1. g_k, x_k may not be jointly w.s.s.

2. estimator filter is turned on at time $k = 0$ with initial conditions (finite memory) and is not LTI.

Objective: find linear min MSE estimate of g_k based only on *finite* past+present measurements

$$\hat{g}_k = \sum_{j=0}^{k} h(k,j) x_j \tag{71}$$

We know that optimal h satisfies the $k \times k$ system of Wiener-Hopf equations.

$$0 = r_{gx}(k,i) - \sum_{j=0}^{k} h(k,j) r_x(j,i), \quad 0 \le i \le k$$

Or, since summation is over finite number of indices we can express this in the familiar matrix form

$$\underline{h}_k = \mathbf{R}_x^{-1} \underline{r}_{xg}$$

where $\underline{h}_k = [h(k,k), h(k,k-1), \ldots, h(k,1)]^T$, $\mathbf{R}_x$ is the $(k+1) \times (k+1)$ covariance matrix of the first k measurements, and $\underline{r}_{gx}$ is the $(k+1)$-element vector of cross correlations between g and $x(0), \ldots, x(k)$.

Difficulty: standard matrix inverse approach has growing memory and computation as k increases: not suitable for real time implementation.

6.6.1 SPECIAL CASE OF UNCORRELATED MEASUREMENTS

As before we first convert to a case where solution to WH is simple:

$\Rightarrow x_i = \eta_i =$ non-stationary white noise

$$r_\eta(j,i) = \sigma_\eta^2(i)\ \delta_{j-i}$$

Solution to WH equation is now immediate

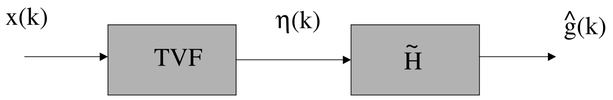


Figure 56: *Decomposition of min MSE filter by prefiltering with time-varying innovations filter.*

$$0 = r_{g\eta}(k,i) - \sum_{j=0}^{k} h(k,j) r_\eta(j,i) = r_{g\eta}(k,i) - h(k,i)\sigma_\eta^2(i), \quad 0 \le i \le k$$

and gives optimal filter as a projection coefficient associeted with projecting g_k onto the i-th noise component η_i

$$\begin{aligned} h(k,i) &= \frac{< g_k, \eta_i >}{< \eta_i, \eta_i >} \\ &= r_{g\eta}(k,i)/\sigma_\eta^2(i), \quad 0 \le i \le k \end{aligned}$$

6.6.2 CORRELATED MEASUREMENTS: THE INNOVATIONS FILTER

Q. How to "prewhiten" x_k?

A. A time-varying "prewhitening filter" has to yield output variables $\{\eta_i\}$ which are uncorrelated, which are causally and linearly generated from past of $\{x_i\}$, and from which the past $\{x_i\}$ can be recovered in a causal fashion

This translates to the following required conditions on η_i:

1. $\text{cov}(\eta_i, \eta_j) = 0$, $i \neq j$
2. $\text{span}\{\eta_k, \eta_{k-1}, \ldots, \eta_0\} = \text{span}\{x_k, x_{k-1}, \ldots, x_0\}$, $k = 1, 2, \ldots$

Recursive construction of $\{\eta_i\}$:

Let $\hat{x}_{k|k-1}$ be the optimal 1-step linear predictor of x_k given past $\{x_{k-1}, \ldots, x_0\}$

$$\hat{x}_{k|k-1} = \sum_{i=0}^{k-1} a_{k,i}\, x_i$$

Equivalently,

$$\hat{x}_{k|k-1} = \sum_{i=0}^{k-1} \alpha_{k,i}\, \eta_i$$

Recall orthogonality condition

$$E[(x_k - \hat{x}_{k|k-1})x_i] = 0, \quad i = 0, \ldots, k-1$$

Suggests following algorithm for η_i's

$$\begin{array}{rcl} \eta_0 &=& x_0 \\ \eta_1 &=& x_1 - \hat{x}_{1|0} \\ \vdots && \vdots \\ \eta_k &=& x_k - \hat{x}_{k|k-1} \end{array}$$

or, more explicitly, in matrix form

$$\begin{bmatrix} \eta_0 \\ \vdots \\ \eta_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -a_{10} & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 \\ -a_{k,0} & \cdots & -a_{k,k-1} & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ \vdots \\ x_k \end{bmatrix}$$

$$\underline{\eta}_k = \mathcal{A}\, \underline{x}_k$$

* η_i is the "innovations process"
* Rows of $\mathcal{A}$ specify causal invertible "innovations filter"
* Note: as $\mathcal{A}$ is invertible $\{\eta_i\}$ is equivalent to $\{x_i\}$.

6.6.3 INNOVATIONS AND CHOLESKY DECOMPOSITION

The innovations representation gives a decomposition for covariance $\mathbf{R}_x$ of $\underline{x}$

$$\begin{array}{rcl} \mathbf{R}_x &=& E[\underline{x}\underline{x}^T] \\ &=& E[\mathcal{A}^{-1}\underline{\eta}\underline{\eta}^T\mathcal{A}^{-T}] \\ &=& \mathcal{A}^{-1}E[\underline{\eta}\underline{\eta}^T]\mathcal{A}^{-T} \\ &=& \mathcal{A}^{-1}\, \mathbf{R}_\eta\, \mathcal{A}^{-T} \end{array}$$

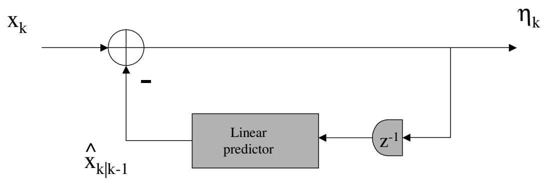


Figure 57: *The innovations filter produces equivalent uncorrelated measurement sequence* η_i.

Equivalently, the representation gives us a way to diagonalize $\mathbf{R}_X$ without the need to form an eigendecomposition:

$$\mathcal{A}\,\mathbf{R}_X\,\mathcal{A}^T = \mathbf{R}_\eta.$$

This decomposition is closely related to the Cholesky decomposition of positive definite matrices [19], which exists in two forms:

FORWARD CHOLESKY DECOMPOSITION

Any symmetric positive definite matrix B has a decomposition of the form

$$\mathbf{B} = \mathbf{L}_f\ \mathbf{P}_f\ \mathbf{L}_f^T$$

where

* $\mathbf{P}_f$ diagonal matrix of "forward prediction error variances"

* $\mathbf{L}_f$ is lower triangular matrix of "forward prediction coefficients"

* $\mathbf{P}_f$ and L_f are non-singular

BACKWARD CHOLESKY DECOMPOSITION

Any symmetric positive definite matrix B has a decomposition of the form

$$\mathbf{B} = \mathbf{L}_b^T\ \mathbf{P}_b\ \mathbf{L}_b$$

where

* $\mathbf{P}_b$ diagonal matrix of "backwards prediction error variances"

* $\mathbf{L}_b$ is lower triangular matrix of "backwards prediction coefficients"

* $\mathbf{P}_b$ and L_b are non-singular

When the measurement sequence $\{x_i\}$ is w.s.s. the covariance matrix $\mathbf{R}_X$ is Toeplitz and there exist fast algorithms, known as the Levinson-Durbin algorithm [19], for diagonalizing the covariance matrix and computing these decompositions.

6.7 TIME VARYING ESTIMATION/PREDICTION VIA THE KALMAN FILTER

Since $\text{span}\{\eta_i\}_{i=0}^k = \text{span}\{x_i\}_{i=0}^k$, the innovations are just another basis spanning the observations, which happens to be a orthogonal basis. Thus we have the equivalent representation for the optimal causal finite memory estimator (71) of g_k

$$\hat{g}_k = \hat{g}_{k|k} = \sum_{j=0}^{k} \tilde{h}(k,j)\eta_j$$

where $\tilde{h}$ is the projection coefficient

$$\tilde{h}(k,j) = \frac{< g_k, \eta_j >}{< \eta_j, \eta_j >} = \frac{r_{g\eta}(k,j)}{\sigma_\eta^2(j)}$$

We can now write a "pseudo recursion" for $\hat{g}_{k|k}$

$$\begin{aligned} \hat{g}_{k|k} &= \sum_{j=0}^{k-1} \tilde{h}(k,j)\eta_j + \tilde{h}(k,k)\ \eta_k \\ &= \hat{g}_{k|k-1} + \tilde{h}(k,k)\ \eta_k \end{aligned} \tag{72}$$

This is not a "true recursion" since we do not yet know how to compute the update $\tilde{h}(k-1,j) \to \tilde{h}(k,j)$ for the projection coefficient nor the update $\eta_{k-1} \to \eta_k$ of the innovations. To obtain a true recursion we will need to assume a dynamical model for x_k. The derivation will then proceed in two steps: first we will consider generating a recursion for the innovations, which will require developing the recursion $\hat{x}_{k|k-1} \to \hat{x}_{k+1|k}$; second we will specialize g_k to the case of signal prediction, $g_k = s_{k+1}$, and signal filtering, $g_k = s_k$, for the case that x_k satisfies the linear model $x_k = s_k + v_k$.

6.7.1 DYNAMICAL MODEL

Specifically, we will assume that x_k is given by the model

$$\begin{aligned} &x_k = s_k + v_k \\ &s_k = \underline{c}_k^T \underline{\xi}_k \\ &\underline{\xi}_{k+1} = \mathbf{A}_{k+1,k}\ \underline{\xi}_k + \mathbf{B}_k\ \underline{w}_k, \quad \underline{\xi}_0 = \underline{\xi}_o \end{aligned} \tag{73}$$

where:

s_k is a (scalar) signal and v_k is a (scalar) measurement noise,

$\underline{\xi}_k$ is a p dimensional "state vector" ("internal state" of system generating $\{\underline{s}_k\}$),

$\underline{c}_k$ is a p-element vector describing how the state affects the signal component s_k of the measurement x_k,

$\underline{w}_k$ is a state noise vector $(q \times 1)$

$\mathbf{B}_k$ is the state noise input matrix $(p \times q)$

$\mathbf{A}_{k+1,k}$ is a state transition matrix $(p \times p)$ describing the signal dynamics which are due soley to the initial condition $\underline{\xi}_o$ (in the absence of driving noise $\underline{w}_k$).

We make the following simplifying statistical assumptions:

State Model (SM) Assumptions A1-A4

- A1 v_k: uncorrelated sequence with zero mean and variance $E[v_k^2] = \sigma_v^2(k)$
- A2 $\underline{w}_k$: uncorrelated sequence with zero mean and covariance matrix $E[\underline{w}_k \underline{w}_k^T] = \mathbf{R}_w(k)$ $(q \times q)$
- A3 $\underline{\xi}_o$ has zero mean and covariance matrix $E[\underline{\xi}_o \underline{\xi}_o^T] = \mathbf{R}_\xi(0)$ $(p \times p)$
- A4 $v_k, \underline{\xi}_0, \underline{w}_j$ are mutually uncorrelated for all j, k

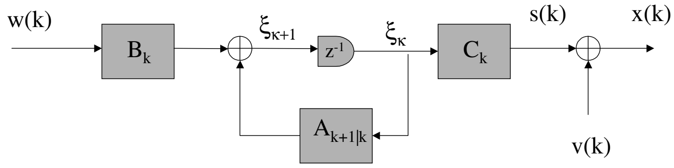


Figure 58: *State space model for observation.*

6.7.2 KALMAN FILTER: ALGORITHM DEFINITION

Here we summarize the Kalman filter as derived in the next section. Under the assumptions A1-A4 the innovations η_k from $\{x_j\}_{j=0}^{k-1}$ can be recursively generated from the following *Kalman recursions for innovation*

Kalman Recursions for innovations and state prediction

First we have the **Measurement Update Equations**:

$$\eta_k = x_k - \underline{c}_k^T \, \hat{\underline{\xi}}_{k|k-1} \tag{74}$$

$$\hat{\underline{\xi}}_{k+1|k} = \mathbf{A}_{k+1,k} \, \hat{\underline{\xi}}_{k|k-1} + \underline{\Gamma}_{k+1,k} \, \eta_k, \quad \hat{\underline{\xi}}_{0|-1} = 0 \tag{75}$$

where $\underline{\Gamma}_{k+1,k}$ is the Kalman Gain, computed offline as function of state predictor error covariance

$$\mathbf{R}_{\tilde{\xi}}(k|k-1) = E[(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k-1})(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k-1})^T]$$

$$\underline{\Gamma}_{k+1,k} \;=\; \mathbf{A}_{k+1,k}\mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k \,\frac{1}{\sigma_\eta^2(k)} \tag{76}$$

where $\sigma_\eta^2(k)$ is the innovations variance

$$\sigma_\eta^2(k) \;=\; \underline{c}_k^T\mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k + \sigma_v^2(k). \tag{77}$$

The covariance matrix $\mathbf{R}_{\tilde{\xi}}(k|k-1)$ is updated according to the **Time Update Equations**:

$$\begin{aligned}\mathbf{R}_{\tilde{\xi}}(k+1|k) &= [\mathbf{A}_{k+1,k} - \underline{\Gamma}_{k+1,k}\underline{c}_k^T]\mathbf{R}_{\tilde{\xi}}(k|k-1)[\mathbf{A}_{k+1,k} - \underline{\Gamma}_{k+1,k}\underline{c}_k]^T \qquad (78)\\ &\quad +\mathbf{B}_k\mathbf{R}_w(k)\mathbf{B}_k^T + \underline{\Gamma}_{k+1,k}\underline{\Gamma}_{k+1,k}^T\sigma_v^2(k).\\ \mathbf{R}_{\tilde{\xi}}(0|-1) &= \mathbf{R}_\xi(0) \qquad (79)\end{aligned}$$

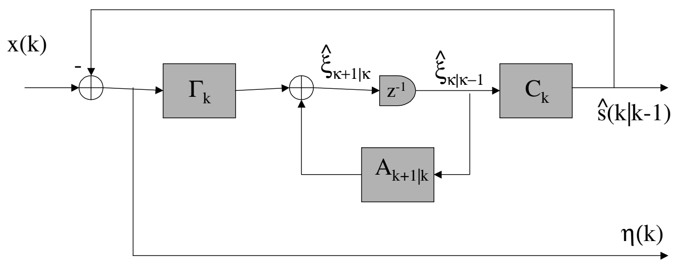


Figure 59: *Kalman filter block diagram for generation of η_k and $\hat{s}(k)$*

6.7.3 KALMAN FILTER: DERIVATIONS

We will derive the Kalman filter in two different ways. The first, called the classical derivation, imposes no distributional assumptions on the state or measurements. The second imposes an additional assumption of Gaussian distributed state and observations. The first derivation historically precedes the second and is based on the projection theorem applied to the innovations process. The second derivation, called the Bayesian derivation, is more direct but less intuitive, relying on posterior density update equations and their differentials.

CLASSICAL KALMAN FILTER DERIVATION

This derivation is based on the Gevers and Kailath innovations approach [18].

Step 1: Gather Together Key Properties

The simplifying assumptions (A1-A4) imply the following properties

Properties P1-P3

P1 $\{\underline{\xi}_j\}_{j=1}^k$ is uncorrelated with $\underline{w}_k$ (A2,A4) and with v_k (A1,A4), and therefore

P2 $\{x_j\}_{j=1}^k$ is uncorrelated with $\underline{w}_k$ and v_{k+1}.

P3 $< x_k, \eta_j > = < s_k + v_k, \eta_j > = < s_k, \eta_j >$, $j < k$, since $< v_k, \eta_j > = 0$, $j < k$, as $\eta_j \in \text{span}\{x_j, x_{j-1}, \ldots, x_0\}$ and v_k is uncorrelated sequence (A1).

Step 2: Establish relation between $\hat{x}_{k|k-1}$ and $\hat{\underline{\xi}}_{k|k-1}$

Putting P2 and P3 together we obtain from representation (72) for $g_k = s_k$:

$$
\begin{aligned}
\hat{x}_{k|k-1} &= \sum_{j=0}^{k-1} \frac{< s_k, \eta_j >}{\|\eta_j\|^2}\eta_j \\
&= \underline{c}_k^T \underbrace{\sum_{j=0}^{k-1} \frac{< \underline{\xi}_k, \eta_j >}{\|\eta_j\|^2}\eta_j}_{\hat{\underline{\xi}}_{k|k-1}} .
\end{aligned}
$$

Step 3: Establish update formula for $\hat{\underline{\xi}}_{k|k-1} \to \hat{\underline{\xi}}_{k+1|k}$

Recall that the linear minimum mean square estimator for a random vector is simply the concatenation of the linear minimum mean square estimators for each element of the random vector. Thus, with abuse of notation, denoting $< \underline{\xi}_{k+1}, \eta_k >$ the vector composed of inner products $< (\xi_{k+1})_i, \eta_k >$, $i = 1, \ldots, p$,

$$
\begin{aligned}
\hat{\underline{\xi}}_{k+1|k} &= \sum_{j=0}^{k} \frac{< \underline{\xi}_{k+1}, \eta_j >}{\|\eta_j\|^2}\eta_j \\
&= \sum_{j=0}^{k-1} \frac{< \underline{\xi}_{k+1}, \eta_j >}{\|\eta_j\|^2}\eta_j + \frac{< \underline{\xi}_{k+1}, \eta_k >}{\|\eta_k\|^2}\eta_k \qquad (80)
\end{aligned}
$$

Define the *Kalman gain* vector

$$
\underline{\Gamma}_{k+1,k} = \frac{< \underline{\xi}_{k+1}, \eta_k >}{\|\eta_k\|^2}, \qquad (81)
$$

and note that, from the state equation (73) for $\underline{\xi}_k$ and the fact that $\underline{w}_k$ and η_j are uncorrelated for $j \leq k$,

$$
\begin{aligned}
\frac{< \underline{\xi}_{k+1}, \eta_j >}{\|\eta_j\|^2} &= \frac{< \mathbf{A}_{k+1,k}\underline{\xi}_k, \eta_j >}{\|\eta_j\|^2} + \frac{< \mathbf{B}_k \underline{w}_k, \eta_j >}{\|\eta_j\|^2} \\
&= \mathbf{A}_{k+1,k}\frac{< \underline{\xi}_k, \eta_j >}{\|\eta_j\|^2}, \quad j \leq k
\end{aligned}
$$

which is $\mathbf{A}_{k+1,k}$ times the projection coefficient for projecting $\underline{\xi}_k$ onto η_j.

Substitution of the above back into (80) gives the desired recursion

$$
\begin{aligned}
\hat{\underline{\xi}}_{k+1|k} &= \mathbf{A}_{k+1,k}\sum_{j=0}^{k-1} \frac{< \underline{\xi}_k, \eta_j >}{\|\eta_j\|^2}\eta_j + \underline{\Gamma}_{k+1,k}\eta_k \\
&= \mathbf{A}_{k+1,k}\hat{\underline{\xi}}_{k|k-1} + \underline{\Gamma}_{k+1,k}\eta_k \qquad (82)
\end{aligned}
$$

with initial condition $\hat{\underline{\xi}}_{0|-1} = 0$.

Step 4: Find expression for $\|\eta_k\|^2$

Define the state estimator error vector

$$\tilde{\underline{\xi}}_{k|k-1} = \underline{\xi}_k - \hat{\underline{\xi}}_{k|k-1}.$$

Then, again using the state model for x_k,

$$\eta_k = x_k - \hat{x}_{k|k-1} = \underline{c}_k^T \underline{\xi}_k + v_k - \underline{c}_k^T \hat{\underline{\xi}}_{k|k-1} = \underline{c}_k^T \tilde{\underline{\xi}}_{k|k-1} + v_k \tag{83}$$

We can use this result to find the innovations variance $\|\eta_k\|^2 = \sigma_\eta^2(k)$ which is required for computing the projection coefficients in (80), specifically the Kalman gain (81) needed for recursion (82). As $\tilde{\underline{\xi}}_{k|k-1} \in \text{span}\{\xi_k, x_{k-1}, \ldots, x_0\}$, $\tilde{\underline{\xi}}_{k|k-1}$ is uncorrelated with v_k, from (83)

$$\sigma_\eta^2(k) = \underline{c}_k^T \mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k + \sigma_v^2(k) \tag{84}$$

where $\mathbf{R}_{\tilde{\xi}}(k|k-1)$ is the state estimator error covariance matrix. To evaluate this we will need to establish a recursion for this error covariance matrix. However, an expression for the Kalman gain will be required to develop the error covariance update equations.

Step 5: Express Kalman gain $\underline{\Gamma}_{k+1,k}$ in terms of state estimator error covariance $\mathbf{R}_{\tilde{\xi}}(k|k-1)$

The Kalman gain vector $\underline{\Gamma}_{k+1,k}$ (81) can be related to the state estimator error covariance by the following steps

1. $<\underline{\xi}_{k+1}, \eta_k> = \mathbf{A}_{k+1,k} <\underline{\xi}_k, \eta_k> + \mathbf{B}_k <\underline{w}_k, \eta_k> = \mathbf{A}_{k+1,k} <\underline{\xi}_k, \eta_k>$ (from whiteness of $\underline{w}_k$ and fact that $\eta_k \in \text{span}\{x_k, \ldots, x_0\}$).
2. $<\underline{\xi}_k, \eta_k> = <\underline{\xi}_k - \hat{\underline{\xi}}_{k|k-1}, \eta_k> = <\tilde{\underline{\xi}}_{k|k-1}, \eta_k>$ (noting that $<\hat{\underline{\xi}}_{k|k-1}, \eta_k> = 0$ from orthogonality principle of linear estimation)
3. $<\tilde{\underline{\xi}}_{k|k-1}, \eta_k> = E[\tilde{\underline{\xi}}_{k|k-1}\eta_k] = E[\tilde{\underline{\xi}}_{k|k-1}\tilde{\underline{\xi}}_{k|k-1}^T]\underline{c}_k$ ($\eta_k = \underline{c}_k^T \tilde{\underline{\xi}}_{k|k-1} + v_k$ and v_k is white noise uncorrelated with $\tilde{\underline{\xi}}_{k|k-1}$)

Putting the above together, and recalling that $\|\eta_k\|^2 = \sigma_\eta^2(k)$ calculated in (84), we obtain

$$\begin{aligned}
\underline{\Gamma}_{k+1,k} &= \frac{<\underline{\xi}_{k+1}, \eta_k>}{\|\eta_k\|^2} = \mathbf{A}_{k+1,k}\frac{<\underline{\xi}_k, \eta_k>}{\|\eta_k\|^2} \\
&= \mathbf{A}_{k+1,k}\mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k \, \frac{1}{\underline{c}_k^T \mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k + \sigma_v^2(k)}
\end{aligned} \tag{85}$$

Step 6: Find recursive update for estimator error covariance $\mathbf{R}_{\tilde{\xi}}(k|k-1) \rightarrow \mathbf{R}_{\tilde{\xi}}(k+1|k)$

First find update equation for $\tilde{\underline{\xi}}_{k|k-1}$ by subtracting the state estimator update equation (82) from the actual state update equation (73)

$$\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k} = \mathbf{A}_{k+1,k}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k-1}) + \mathbf{B}_k \underline{w}_k - \underline{\Gamma}_{k+1,k}\eta_k.$$

Identifying $\tilde{\underline{\xi}}_{k+1|k} = \underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k}$ and using $\eta_k = \underline{c}_k^T \tilde{\underline{\xi}}_{k|k-1} + v_k$ in the above

$$\begin{aligned} \tilde{\underline{\xi}}_{k+1|k} &= \mathbf{A}_{k+1,k}\tilde{\underline{\xi}}_{k|k-1} + \mathbf{B}_k \underline{w}_k - \underline{\Gamma}_{k+1,k}(\underline{c}_k^T \tilde{\underline{\xi}}_{k|k-1} + v_k) \\ &= [\mathbf{A}_{k+1,k} - \underline{\Gamma}_{k+1,k}\underline{c}_k^T]\tilde{\underline{\xi}}_{k|k-1} + \mathbf{B}_k \underline{w}_k - \underline{\Gamma}_{k+1,k} v_k \end{aligned} \tag{86}$$

Now properties P1-P3 imply that the three additive terms in (86) are mutually uncorrelated. Therefore, the covariance of the RHS is the sum of three covariance matrices and we obtain the update equation (79)

$$\begin{aligned} \mathbf{R}_{\tilde{\xi}}(k+1|k) &= [\mathbf{A}_{k+1,k} - \underline{\Gamma}_{k+1,k}\underline{c}_k^T]\mathbf{R}_{\tilde{\xi}}(k|k-1)[\mathbf{A}_{k+1,k} - \underline{\Gamma}_{k+1,k}\underline{c}_k^T]^T \\ &\quad + \mathbf{B}_k \mathbf{R}_w(k)\mathbf{B}_k^T + \underline{\Gamma}_{k+1,k}\underline{\Gamma}_{k+1,k}^T \sigma_v^2(k). \end{aligned}$$

An alternative form of the recursion (79) can be derived by using (77) and (81) which makes $\mathbf{R}_{\tilde{\xi}}(k|k-1)$ explicit in the quantity $\underline{\Gamma}_{k+1,k}$. After some algebra this produces the equivalent update equation

$$\begin{aligned} \mathbf{R}_{\tilde{\xi}}(k+1|k) &= \mathbf{A}_{k+1,k}\mathbf{R}_{\tilde{\xi}}(k|k-1)\mathbf{A}_{k+1,k}^T + \mathbf{B}_k \mathbf{R}_w(k)\mathbf{B}_k^T \\ &\quad - \mathbf{A}_{k+1,k}\mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k^T \underline{c}_k \mathbf{R}_{\tilde{\xi}}(k|k-1)\mathbf{A}_{k+1,k}^T / \sigma_\eta^2(k) \end{aligned} \tag{87}$$

where $\sigma_\eta^2(k) = \underline{c}_k^T \mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k + \sigma_v^2(k)$, as defined above.

Drawing all of the results of this subsection together we obtain the Kalman filtering equations (74)-(79) of Section 6.7.2.

BAYESIAN KALMAN FILTER DERIVATION

Similarly to Section 6.7.3 we assume that the observations y_k obey a dynamical model, except that here we also assume that the additive noises and the initial state are *Gaussian*. Specifically,

$$\begin{aligned} y_k &= s_k + v_k, \\ s_k &= \underline{c}^T \underline{\xi}_k, \qquad\qquad k = 0, 1, \ldots \\ \underline{\xi}_{k+1} &= \mathbf{A}\underline{\xi}_k + \mathbf{B}_k \underline{w}_k, \end{aligned}$$

where $\underline{\xi}_0$, v_k and $\underline{w}_k$ are mutually independent zero mean temporally uncorrelated Gaussian variables for $k = 0, 1, \ldots$. As before v_k and $\underline{w}_k$ are zero mean (white) noises with variance σ_v^2 and covariance matrix $\mathbf{R}_w$, respectively. All other assumptions on the model are identical to those made in the previous section. Let the posterior density of the state $\underline{\xi}_k$ given the observation sequence $\mathcal{Y}_l = \{y_l, y_{l-1}, \ldots, y_0\}$ up to time l be denoted $f_{\underline{\xi}_k|\mathcal{Y}_l}$.

The Bayesian derivation of the Kalman filter equations is split into 4 steps.

Step 1: Show that posterior density of state is a multivariate Gaussian density

The key to the Bayes derivation of the Kalman filter is that the posterior density $f_{\underline{\xi}_k|\mathcal{Y}_k}$ must be of the form

$$f_{\underline{\xi}_k|\mathcal{Y}_l}(\underline{\xi}_k|\mathcal{Y}_k) = \frac{1}{|\mathbf{R}_{\tilde{\xi}}(k|k)|^{\frac{1}{2}}(2\pi)^{p/2}} \exp\left(-\frac{1}{2}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})^T \mathbf{R}_{\tilde{\xi}}^{-1}(k|k)(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})\right) \tag{88}$$

where $\hat{\underline{\xi}}_{k|k}$ is the Kalman filter's state estimator and $\mathbf{R}_{\tilde{\xi}}(k|k) = E[(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})^T]$ is the state estimator error covariance matrix.

To see that relation (88) holds, recall that for any two jointly Gaussian r.v.s W, Z the conditional distribution of Z given W is $\mathcal{N}(E[Z|W], \text{cov}(Z|W))$. $E[Z|W]$ is identical to the linear minimum mean squared error estimator of Z given W, which, for $Z = \underline{\xi}_k$ and $W = \mathcal{Y}_k$ must be the output of the Kalman filter. Furthermore, the conditional covariance $\text{cov}(\underline{\xi}_k - \hat{\underline{\xi}}_k|\mathcal{Y}_k) = \text{cov}(\underline{\xi}_k - \hat{\underline{\xi}}_k)$ since (projection theorem) the error $\underline{\xi}_k - \hat{\underline{\xi}}_k$ is uncorrelated with past observations $\mathcal{Y}_k$ and therefore, since the error and the past observations are jointly Gaussian, they are statistically independent. Hence the conditional covariance is equal to the unconditional error covariance $R_{\tilde{\xi}}(k|k)$ of the Kalman state estimator.

Step 2: Derive update equation for posterior density of state

The next step is to show that the state posterior density $\rho_k(\underline{\xi}_k) \stackrel{\text{def}}{=} f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_k|\mathcal{Y}_k)$ obeys the *Chapman-Kolmogorov* formula

$$\rho_{k+1}(\underline{\xi}_{k+1}) = \frac{f_{y_{k+1}|\underline{\xi}_{k+1}}(y_{k+1}|\underline{\xi}_{k+1})}{f_{y_{k+1}|\mathcal{Y}_k}(y_{k+1}|\mathcal{Y}_k)} \int_{\mathbb{R}^p} f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k)\rho_k(\underline{\xi}_k)d\underline{\xi}_k.. \tag{89}$$

This formula is valid even if $\underline{\xi}_o$, v_k, $\underline{w}_k$ are not Gaussian.

To show (89) start with Bayes formula

$$\begin{aligned} f_{\underline{\xi}_{k+1}|\mathcal{Y}_{k+1}}(\underline{\xi}_{k+1}|\mathcal{Y}_{k+1}) &= f_{\underline{\xi}_{k+1}|\mathcal{Y}_{k+1}}(\underline{\xi}_{k+1}|y_{k+1}, \mathcal{Y}_k) \\ &= \frac{f_{\underline{\xi}_{k+1}, y_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1}, y_{k+1}|\mathcal{Y}_k)}{f_{y_{k+1}|\mathcal{Y}_k}(y_{k+1}|\mathcal{Y}_k)} \end{aligned}$$

Next we express the numerator as

$$\begin{aligned} f_{\underline{\xi}_{k+1}, y_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1}, y_{k+1}|\mathcal{Y}_k) &= f_{y_{k+1}|\underline{\xi}_{k+1}, \mathcal{Y}_k}(y_{k+1}|\underline{\xi}_{k+1}, \mathcal{Y}_k) f_{\underline{\xi}_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1}|\mathcal{Y}_k) \\ &= f_{y_{k+1}|\underline{\xi}_{k+1}}(y_{k+1}|\underline{\xi}_{k+1}) f_{\underline{\xi}_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1}|\mathcal{Y}_k) \end{aligned}$$

where in the last step we used the fact that, given $\underline{\xi}_{k+1}$, y_{k+1} is independent of the past $\mathcal{Y}_k$ since the noise v_k is white (recall the model $y_{k+1} = \underline{c}^T\underline{\xi}_{k+1} + v_k$). This establishes the Chapman-Kolmogorov recursive formula (89)

When $\underline{\xi}_o$, v_k, $\underline{w}_k$ are Gaussian the density $f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{u}|\underline{\xi}_k)$ has the form of a multivariate Gaussian density over $\underline{u}$ with mean parameter $\underline{g}(\underline{\xi}_k)$ and covariance matrix parameter $\mathbf{B}\mathbf{R}_w\mathbf{B}^T$ and the density $f_{y_{k+1}|\underline{\xi}_{k+1}}(z|\underline{\xi}_{k+1})$ has the form of a univariate Gaussian density over z with mean parameter $\underline{c}^T\underline{\xi}_{k+1}$ and variance parameter σ_v^2. Indeed, given $\underline{\xi}_{k+1}$, y_{k+1} is obviously Gaussian distributed with mean $\underline{c}^T\underline{\xi}_{k+1}$ and variance σ_v^2. To show the Gaussian form of $f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{u}|\underline{\xi}_k)$ use the *law of total probability* to obtain

$$\begin{aligned} f_{\underline{\xi}_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1}|\mathcal{Y}_k) &= \int_{\mathbb{R}^p} f_{\underline{\xi}_{k+1}, \underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_{k+1}, \underline{\xi}_k|\mathcal{Y}_k)d\underline{\xi}_k \\ &= \int_{\mathbb{R}^p} f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_k|\mathcal{Y}_k) f_{\underline{\xi}_{k+1}|\underline{\xi}_k, \mathcal{Y}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k, \mathcal{Y}_k)d\underline{\xi}_k. \end{aligned}$$

Now, as $\underline{\xi}_{k+1}$ is independent of $\mathcal{Y}_k$ given $\underline{\xi}_k$ (recall that $\underline{w}_k$ and v_k are independent Gaussian white noises), the second factor in the integrand is simply $f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k)$ which is a multivariate Gaussian density.

Step 3: Find expression for exponents in posterior density update equation

Next we derive a relation for the quadratic form $(\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1})^T \mathbf{R}_{\underline{\xi}}^{-1}(k+1|k+1)(\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1})$ by equating the exponent on the left hand side of the Chapman-Kolmogorov equations to the exponent on the right hand side using the Gaussian forms of all the densities expressed in these equations.

Completion of the square in the integrand of (89) gives the expression

$$
\begin{aligned}
&f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k)\rho_k(\underline{\xi}_k) \qquad (90)\\
&= c\exp\left(-\frac{1}{2}(\underline{\xi}_{k+1} - \mathbf{A}\hat{\underline{\xi}}_{k|k} - \mathbf{A}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k}))^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}(\underline{\xi}_{k+1} - \mathbf{A}\hat{\underline{\xi}}_{k|k} - \mathbf{A}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k}))\right)\\
&\exp\left(-\frac{1}{2}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})^T \mathbf{R}_{\underline{\xi}}^{-1}(k|k)(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})\right)\\
&= c\exp\left(-\frac{1}{2}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k} - \mathbf{Q}^{-1}\underline{u}_k)^T\mathbf{Q}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k} - \mathbf{Q}^{-1}\underline{u}_k)\right)\exp\left(-\frac{1}{2}(q_1 - q_2)\right)
\end{aligned}
$$

where c is an unimportant constant and

$$
\begin{aligned}
\underline{u} &= [\mathbf{A}^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}\mathbf{A}(\underline{\xi}_{k+1} - \mathbf{A}\hat{\underline{\xi}}_{k|k})\\
\mathbf{Q} &= \mathbf{A}^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}\mathbf{A} + \mathbf{R}_{\underline{\xi}}^{-1}(k|k)\\
q_1(\underline{\xi}_{k+1}) &= (\underline{\xi}_{k+1} - \mathbf{A}\hat{\underline{\xi}}_{k|k})^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}(\underline{\xi}_{k+1} - \mathbf{A}\hat{\underline{\xi}}_{k|k})\\
q_2(\underline{\xi}_{k+1}) &= \underline{u}^T\mathbf{Q}^{-1}\underline{u}.
\end{aligned}
$$

The result of integration of () over $\underline{\xi}_k \in \mathbb{R}^p$ (recall that the Gaussian density integrates to 1) gives the following expression for the integral in (89)

$$
c_1 \exp\left(-\frac{1}{2}(q_1 - q_2)\right)
$$

for some constant c_1. Now the exponent on the RHS of (89) can be easily found

$$
-\frac{1}{2}\left((y_{k+1} - \underline{c}^T\underline{\xi}_{k+1})^2/\sigma_v^2 - (y_{k+1} - \underline{c}^T\hat{\underline{\xi}}_{k+1|k+1})^2/\sigma^2 + q_1(\underline{\xi}_{k+1}) + -q_2(\underline{\xi}_{k+1})\right) \qquad (91)
$$

where $\sigma^2 = \text{var}(y_{k+1} - \underline{c}^T\hat{\underline{\xi}}_{k+1|k+1}) = \underline{c}^T\mathbf{R}_{\underline{\xi}}(k+1|k+1)\underline{c} + \sigma_v^2$. Thus we have the relation

$$
\begin{aligned}
&(\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1})^T\mathbf{R}_{\underline{\xi}}^{-1}(k+1|k+1)(\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1}) = \qquad (92)\\
&\quad (y_{k+1} - \underline{c}^T\underline{\xi}_{k+1})^2/\sigma_v^2 - (y_{k+1} - \underline{c}^T\hat{\underline{\xi}}_{k+1|k+1})^2/\sigma^2 + q_1(\underline{\xi}_{k+1}) + -q_2(\underline{\xi}_{k+1})
\end{aligned}
$$

Step 3: Differentiate the exponent of the posterior update equation

Using the relation (92) we now derive the Kalman filter equations specifying state estimator updates $\hat{\underline{\xi}}_{k|k} \rightarrow \hat{\underline{\xi}}_{k+1|k+1}$ and inverse covariance updates $\mathbf{R}_{\underline{\xi}}^{-1}(k|k) \rightarrow \mathbf{R}_{\underline{\xi}}^{-1}(k+1|k+1)$ in the following manner. To derive state update equation we take the derivative of relation (92) with respect to $\underline{\xi}_{k+1}$ and evaluate the resulting equation at $\underline{\xi}_{k+1} = \mathbf{A}\hat{\underline{\xi}}_{k|k}$. To derive the covariance update equation we take the second derivative with respect to $\underline{\xi}_{k+1}$. Here are the details.

Differentiation of the LHS of (92) twice in the argument $\underline{\xi}_{k+1}$ yields $2\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)$. Likewise twice differentiating the RHS of (92) and equating to $2\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)$ gives

$$\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1) = [\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1} - [\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}\mathbf{A}[\mathbf{A}^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}\mathbf{A} + \mathbf{R}_{\tilde{\xi}}^{-1}(k|k)]^{-1}\mathbf{A}^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1} + \frac{\underline{c}\underline{c}^T}{\sigma_v^2}$$

Application of the Sherman-Morrison-Woodbury identity (1) to the first two terms on the RHS gives a compact recursion for the inverse covariance

$$\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1) = [\mathbf{B}\mathbf{R}_w\mathbf{B}^T + \mathbf{A}\mathbf{R}_{\tilde{\xi}}(k|k)\mathbf{A}^T]^{-1} + \frac{\underline{c}\underline{c}^T}{\sigma_v^2} \tag{93}$$

Next we differentiate the LHS and RHS of (92) once wrt $\underline{\xi}_{k+1}$ and evaluate at $\underline{\xi}_{k+1} = \mathbf{A}\hat{\underline{\xi}}_{k|k}$ to obtain

$$\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)(\mathbf{A}\hat{\underline{\xi}}_{k|k} - \hat{\underline{\xi}}_{k+1|k+1}) = -\frac{\underline{c}}{\sigma_v^2}(y_{k+1} - \underline{c}^T\mathbf{A}\hat{\underline{\xi}}_{k|k})$$

Which yields the Kalman filter recursion

$$\hat{\underline{\xi}}_{k+1|k+1} = \mathbf{A}\hat{\underline{\xi}}_{k|k} + \Gamma_{k|k}(y_{k+1} - \underline{c}^T\mathbf{A}\hat{\underline{\xi}}_{k|k})$$

where we have identified the Kalman gain

$$\Gamma_{k|k} = \mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)\frac{\underline{c}}{\sigma_v^2}.$$

Thus we obtain the Kalman filtering equations (74)-(79) of Section 6.7.2.

6.8 KALMAN FILTERING: SPECIAL CASES

The Kalman filter equation (75) generates the innovations sequence η_k which is needeed to compute the estimate $\hat{g}_{k|k}$ defined in Sec. 6.7 by the equation (72). Also needed are the projection coefficients $\tilde{h}(k,j)$, $j = 1, \ldots, k$. We discuss two special cases for which these coefficients are simply computed, Kalman prediction and Kalman filtering

6.8.1 KALMAN PREDICTION

The linear prediction problem is to predict future value of the observation x_{k+1} from a linear combination of past and present observations $\{x_j\}_{j=0}^k$, or, equivalently, from the past and present innovations $\{\eta_j\}_{j=0}^k$. Recalling the measurement model (73), $x_{k+1} = s_{k+1} + v_{k+1}$ is the sum of two uncorrelated components. Hence, denoting the predictor by $\hat{x}_{k+1|k}$ and applying the superposition property (60) of linear estimators of a sum of random variables

$$\hat{x}_{k+1|k} = \hat{s}_{k+1|k} + \hat{v}_{k+1|k} = \hat{s}_{k+1|k}$$

where $\hat{v}_{k+1|k} = 0$ due to the fact that v_k is white and thus uncorrelated with the past innovations, i.e. unpredictable. Finally, as $s_{k+1} = \underline{c}_{k+1}^T\underline{\xi}_{k+1}$

$$\hat{s}_{k+1|k} = \underline{c}_{k+1}^T\hat{\underline{\xi}}_{k+1|k}$$

which can be computed from the Kalman filter (75) for state estimation discussed in the previous sub-section.

6.8.2 KALMAN FILTERING

The filtering problem is to estimate the signal component s_k in $x_k = s_k + v_k$ from past and present measurements $\{x_j\}_{j=0}^k$ (equivalently $\{\eta_j\}_{j=0}^k$). Let $\hat{s}_{k|k}$ denote this estimate. Set $g_k = s_k$ and from the general recursion (72) we obtain

$$\hat{s}_{k|k} = \hat{s}_{k|k-1} + \tilde{h}(k,k)\eta_k,$$

where $\hat{s}_{k|k-1}$ is the linear predictor derived in the last subsection and

$$\tilde{h}(k,k) = \frac{E[s_k\eta_k]}{\mathrm{var}(\eta_k)} = \frac{\underline{c}_k^T E[\underline{\xi}_k \eta_k]}{\mathrm{var}(\eta_k)}.$$

Recall that in the process of showing the expression (85) for the Kalman gain $\underline{\Gamma}_{k+1,k}$ we established $E[\underline{\xi}_k\eta_k] = \mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k$. Putting this together with the expression (77) we obtain

$$\tilde{h}(k,k) = \frac{\underline{c}_k^T \mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k}{\underline{c}_k^T \mathbf{R}_{\tilde{\xi}}(k|k-1)\underline{c}_k + \sigma_v^2(k)}.$$

All of the above quantites are available from the Kalman filter recursions (74) and (79).

6.9 KALMAN FILTER FOR SPECIAL CASE OF GAUSSIAN STATE AND NOISE

Assume:

* v_k, $\underline{\xi}_o$ and $\underline{w}_k$ are jointly Gaussian

Then:

* $x_k = s_k + v_k$ is a Gaussian random process

* Kalman filter yields min MSE state predictor

$$\hat{\underline{\xi}}_{k|k-1} = E[\underline{\xi}_k | x_{k-1}, \ldots, x_1]$$

* $\{\eta_k\}$ is an equivalent uncorrelated Gaussian measurement

6.10 STEADY STATE KALMAN FILTER AND WIENER FILTER

Assume

* $\mathbf{A}_{k+1,k}, b_k, c_k$ and $\mathbf{R}_w(k)$ are time-invariant

* $\mathbf{R}_v(k)$ is time-invariant

* The state error covariance matrix $\mathbf{R}_{\tilde{\xi}}(k+1,k)$ converges to a positive definite matrix as $k \to \infty$.

Then:

* s_k is w.s.s. as $k \to \infty$

* x_k is w.s.s. as $k \to \infty$

$\Rightarrow$ Steady state innovations filter is equivalent to Wiener prewhitening filter

In particular, in steady state, η_k becomes a (w.s.s.) white noise with

$$\sigma^2_\eta(k) \to \sigma^2_\eta(\infty) = \underline{c}^T \mathbf{R}_{\tilde{\xi}}(\infty)\underline{c} + \sigma^2_v$$

The steady state error covariance matrix $\mathbf{R}_{\tilde{\xi}}(\infty)$ can be found in two steps:

Step 1: set $\mathbf{R}_{\tilde{\xi}}(k, k-1) = \mathbf{R}_{\tilde{\xi}}(k+1, k) = \mathbf{R}_{\tilde{\xi}}(\infty)$ in covariance update equation (79), equivalently, (87), obtaining the steady state covariance equation:

$$\begin{aligned}\mathbf{R}_{\tilde{\xi}}(\infty) &= \mathbf{A}\mathbf{R}_{\tilde{\xi}}(\infty)\mathbf{A}^T + \mathbf{B}\mathbf{R}_w\mathbf{B}^T \\ &\quad -\mathbf{A}\mathbf{R}_{\tilde{\xi}}(\infty)\underline{c}^T\underline{c}\mathbf{R}_{\tilde{\xi}}(\infty)\mathbf{A}^T/\sigma^2_\eta(\infty),\end{aligned}$$

Step 2: Noting that, as $\sigma^2_\eta(\infty)$ is linear in $\mathbf{R}_{\tilde{\xi}}(\infty)$, the steady state covariance equation is equivalent to a quadratic equation in $\mathbf{R}_{\tilde{\xi}}(\infty)$, called an *algebraic Ricatti equation* [32]. This can be solved numerically but for small state dimension $\mathbf{R}_{\tilde{\xi}}(\infty)$ it can be often found by hand.

Example 30 *Kalman filter for estimation of a constant signal*

This example is credited to R. Raich. The objective is to find an optimal recursive estimator of a constant signal in random noise given a finite number of observations. Accordingly, let's assume the following special case of the dynamical observation model (73)

$$\begin{aligned}x_k &= s_k + v_k \\ s_{k+1} &= s_k.\end{aligned}$$

Here s_k is a scalar state and we can identify $\underline{c}_k = 1, \mathbf{B}_k = 0, \mathbf{A}_{k+1,k} = 1$, and $\mathbf{R}_\xi(0) = \sigma^2_s$. For notational simplicity define the normalized state error covariance (actually the variance since the state is one dimensional):

$$T_{k+1} = \mathbf{R}_{\tilde{\xi}}(k+1, k)/\sigma^2_v.$$

With this notation, and the identifications above, the (scalar) update equation (87) for $\mathbf{R}_{\tilde{\xi}}(k+1, k)$ gives

$$T_{k+1} = T_k/(T_k + 1), \quad T_o = \sigma^2_s/\sigma^2_v,$$

which has explicit solution $T_k = 1/(k + 1/\text{SNR})$, where SNR$= \sigma^2_s/\sigma^2_v$. The Kalman gain is simply

$$\Gamma_{k+1,k} = \frac{T_k}{T_k + 1} = T_{k+1}.$$

Therefore, the Kalman filter update for $\hat{s}_{k|k-1}$ is

$$\hat{s}_{k+1|k} = \hat{s}_{k|k-1} + \Gamma_{k+1,k}\eta_k,$$

which, using $\eta_k = x_k - \hat{s}_{k|k-1}$ is equivalent to the AR(1) recursion

$$\hat{s}_{k+1|k} = [1 - \Gamma_{k+1,k}]\hat{s}_{k|k-1} + \Gamma_{k+1,k}x_k,$$

with initial condition $\hat{s}_{0|-1} = 0$. approximation to $\Gamma_{k+1,k} = T_{k+1}$:

$$T_{k+1} \approx \frac{1}{k+1},$$

yielding the large k form of the AR(1) recursion:

$$\hat{s}_{k+1|k} = \frac{k}{k+1}\hat{s}_{k|k-1} + \frac{1}{k+1}x_k,$$

which is equivalent to

$$\hat{s}_{k+1|k} = \frac{1}{k+1}\sum_{i=0}^{k} x_k.$$

Thus, as expected, the Kalman filter estimator of a constant signal becomes identical to the sample mean estimator of the ensemble mean for large k - as the transients of the filter die down the initial condition has no more influence.

It should be observed in the above example that the Kalman filter does not converge in steady state to a LTI filter since the asymptotic state covariance is not positive definite - the variance is equal to zero.

6.11 SUMMARY OF STATISTICAL PROPERTIES OF THE INNOVATIONS

We summarize important properties of the innovations that will be important in the sequel. As the observation noise v_k is uncorrelated with the signal s_k, we have three equivalent expressions for the innovations

$$\begin{aligned}
\eta_k &= x_k - \hat{x}_{k|k-1} \\
\eta_k &= x_k - \hat{s}_{k|k-1} \\
\eta_k &= \underline{c}_k^T(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k-1}) + v_k
\end{aligned}$$

Furthermore:

$$\begin{aligned}
&E[\eta_i] = 0 \\
&\mathrm{cov}(\eta_i, \eta_j) = 0, \quad i \neq j
\end{aligned}$$

and, as shown above, the innovations variance is

$$\begin{aligned}
\mathrm{var}(\eta_k) &= \sigma_\eta^2(k) \\
&= \underline{c}_k^T \mathbf{R}_{\tilde{\xi}}(k)\underline{c}_k + \sigma_v^2(k)
\end{aligned}$$

6.12 BACKGROUND REFERENCES

The Wiener filter was originally published (as a classified report) by Norbert Wiener in the early 1940's and the Kalman filter was published by Kalman and Bucy in the early 1960's. The book by Kailath [32] provides a nice overview of the historical context for both of these breakthroughs.

Other books covering linear prediction from a signal processing perspective are Hayes [23], Mendel [45], and Moon and Stirling [49]. A very comprehensive mathematically concise coverage of signal processing algorithms for non-statistical least squares and linear prediction can be found in the book by Strobach [70]. Finally, for a different time series perspective of mathematical statistics the reader can consult the excellent book by Brockwell and Davis [12].

6.13 APPENDIX: POWER SPECTRAL DENSITIES

Here we provide a quick primer on autocorrelation functions (acf) and power spectral densities (PSD) for zero mean wide sense stationary (wss) random sequences. For more detailed information see Thomas [72] or Davenport and Root [13].

6.13.1 ACF AND CCF

The autocorrelation function of a zero mean finite variance discrete time random process $\{x_k\}$ is defined as

$$r_x(i,j) = E[x_i x_j^*].$$

The acf is non-negative definite in the sense that for any absolutely summable sequence $\{u_k\}$

$$\sum_{i,j} u_i^* r_x(i,j) u_j \geq 0.$$

For two zero mean finite variance discrete time random sequences x_k and y_k the cross-correlation function (ccf) of x and y is defined as

$$r_{xy}(i,j) = E[x_i y_j^*].$$

The ccf has conjugate symmetry, $r_{xy}(i,j) = r_{yx}^*(j,i)$, and is equal to zero when x and y are uncorrelated random sequences.

6.13.2 REAL VALUED WIDE SENSE STATIONARY SEQUENCES

When x_k is zero mean real and wss its acf satisfies (by definition of wss): $r_x(i,j) = r_x(i-j,0)$. The function $r_x(i,i-k)$ is usually denoted as $r_x(k)$ and it is conjugate symmetric

$$r_x(-k) = r_x(k)$$

and satisfies $r_x(0) \geq r_x(k)$ for all k. A real wss x_k has a PSD defined as the Discrete Time Fourier Transform (DTFT) of r_x:

$$\mathcal{P}_x(\omega) = \mathcal{F}\{r_x(k)\} = \sum_{k=-\infty}^{\infty} r_x(k) e^{-j\omega k}$$

from which r_x can be recovered using the inverse DTFT:

$$r_x(k) = \mathcal{F}^{-1}\{\mathcal{P}_x(\omega)\} = \frac{1}{2\pi}\int_{\pi}^{\pi} \mathcal{P}_x(\omega) e^{j\omega k}.$$

Due to the real, symmetric, non-negative definiteness of r_x its PSD is real, symmetric, and non-negative:

$$\mathcal{P}_x^*(\omega) = \mathcal{P}_x(\omega), \quad \mathcal{P}_x(-\omega) = \mathcal{P}_x(\omega), \quad \mathcal{P}_x(\omega) \geq 0.$$

For two zero mean real jointly wss random sequences x_k and y_k the ccf similarly satisfies: $r_{xy}(i, i-k) = r_{xy}(k)$. The ccf also has a kind of symmetry property $r_{xy}(-k) = r_{yx}(k)$. The cross PSD is defined as the DTFT of r_{xy}. It is neither real nor symmetric but inherits the following property from r_{xy}

$$\mathcal{P}_{xy}(-\omega) = \mathcal{P}_{xy}^*(\omega) = \mathcal{P}_{yx}(\omega).$$

When x_k and y_k are uncorrelated the CPSD is equal to zero and therefore if $z_k = x_k + y_k$

$$\mathcal{P}_z(\omega) = \mathcal{P}_x(\omega) + \mathcal{P}_y(\omega).$$

Let h_k be the impulse response of a stable linear time invariant (LTI) system and let $H(\omega) = \mathcal{F}(h_k)$ be its frequency-domain transfer function. If $y_k = h_k * x_k$ is the output of this LTI system then

$$\mathcal{P}_y(\omega) = |H(\omega)|^2 \mathcal{P}_x(\omega),$$

and

$$\mathcal{P}_{xy}(\omega) = H^*(\omega)\mathcal{P}_x(\omega), \;\; \mathcal{P}_{yx}(\omega) = H(\omega)\mathcal{P}_x(\omega).$$

6.13.3 Z-DOMAIN PSD AND CPSD

Analogously to the frequency domain PSD defined above, one can define the z-domain PSD by

$$\tilde{\mathcal{P}}_x(z) = \mathcal{Z}\{r_x(k)\} = \sum_{k=-\infty}^{\infty} r_x(k) z^{-k}$$

We have the obvious relation $\mathcal{P}_x(\omega) = \tilde{\mathcal{P}}_x(e^{j\omega})$. The z-domain CPSD is defined analogously: $\tilde{\mathcal{P}}_{xy}(z) = \mathcal{Z}\{r_{xy}(k)\}$. The z-domain PSD and CPSD inherit various properties from their frequency-domain versions. In particular, radial symmetry about the unit circle

$$\tilde{\mathcal{P}}_x(z^{-1}) = \tilde{\mathcal{P}}_x(z), \;\; \tilde{\mathcal{P}}_{xy}(z^{-1}) = \tilde{\mathcal{P}}_{yx}(z)$$

and conjugate symmetry

$$\tilde{\mathcal{P}}_x(z^*) = \tilde{\mathcal{P}}_x^*(z).$$

Finally, when $y_k = h_k * x_k$ is the output of an LTI system with z-domain transfer function $\tilde{H}(z)$ then we have the analogous relations to before

$$\tilde{\mathcal{P}}_y(z) = \tilde{H}(z)\tilde{H}(z^{-1})\tilde{\mathcal{P}}_x(z),$$

and

$$\tilde{\mathcal{P}}_{xy}(z) = \tilde{H}(z^{-1})\tilde{\mathcal{P}}_x(z), \quad \tilde{\mathcal{P}}_{yx}(z) = \tilde{H}(z)\tilde{\mathcal{P}}_x(z).$$

NB: When there is little danger of confusion, we usually drop the tilde notation when expressing z-domain PSDs or transfer functions

6.14 EXERCISES

6.1 As we know, the optimal filter $h(k,j)$ for estimating the sample $g(k)$ of the process $\{g(k)\}_{i=-\infty}^{\infty}$ from zero mean process $\{x(i)\}_{i=-\infty}^{\infty}$ satisfies the Wiener-Hopf equation

$$r_{gx}(k,i) - \sum_{j=-\infty}^{\infty} h(k,j) r_x(i,j) = 0, \quad -\infty < i < \infty$$

Show that when g and x are jointly w.s.s. random processes, i.e. $r_x(i,j) = r_x(i-j)$ and $r_{gx}(k,i) = r_{gx}(k-i)$, $h(k,j)$ can be assumed to be linear time invariant (LTI), i.e. $h(k,j) = h(k-j)$ will satisfy the WH equation. Now show that the same holds for the causal optimal filter $h(k,j)$ which satisfies

$$r_{gx}(k,i) - \sum_{j=-\infty}^{\infty} h(k,j) r_x(i,j) = 0, \quad -\infty < i \leq k.$$

(Hint: Find Weiner-Hopf equations for estimating $y(k+l)$ where l is an arbitrary time shift, make a change of index i and a change of variable j and show by comparison to the Weiner-Hopf equations for estimating $y(k)$ that $h(k,j) = h(k-j,0)$).

6.2 Derive the equation (65) for the Causal Wiener filter $0 = r_{gx}(l) - \sum_{m=-\infty}^{\infty} h(l-m) r_x(m)$ from the original equation (64) by making two changes of variable in sequence: reindex i by $l = k - i$ and reindex j by $m = k - j$.

6.3 For constants a, c and the definitions $q = 1 + c + ca^2$, $r = ca$, derive the following identity

$$\frac{1}{(1+az^{-1})} \frac{1}{(1+az)} + c = \frac{d(1+bz^{-1})}{(1+az^{-1})} \frac{d(1+bz)}{(1+az)} \tag{94}$$

where

$$b = \frac{q/r \pm \sqrt{(q/r)^2 - 4}}{2}, \qquad d^2 = q/(1+b^2)$$

Observe that when c is positive real, one of the roots in the equation for b satisfies $|b| \leq 1$ while the other satisfies $|b| \geq 1$.

6.4 In the development of the causal and non-causal Weiner estimators we have assumed that all processes were zero mean. Here we deal with the case of non-zero mean w.s.s. processes for which the *affine Weiner estimator* is appropriate.

Assume the measurement process $\{x_k\}_k$ and the target process $\{g_k\}_k$ to be estimated are w.s.s. and have non-zero means $E[x_k] = \mu_x(k)$ and $E[g_k] = \mu_g(k)$.

(a) The affine non-causal Wiener estimator of g_k is defined by the filter $h(k,j)$ and sequence of constants s_k as

$$\hat{g}_k = a_k + \sum_{j=-\infty}^{\infty} h(k,j) x_j$$

where $h(k,j)$ and a_k are selected to minimize the mean square estimation error $\text{MSE}(h,a) = E[(g_k - \hat{g}_k)^2]$. Show that the optimal filter satisfies $h(k,j) = h(k-j)$ where $h(j)$ is the optimal linear time invariant non-causal Weiner filter for estimating the zero mean process $g_k - \mu_g(k)$ from the centered zero mean measurements $x_k - \mu_x(k)$ and that the optimal sequence a_k is $\mu_g(k) + \sum_{j=-\infty}^{\infty} h(k-j)\mu_x(j)$. Further show that the minimum MSE of the affine non-causal Wiener estimator is functionally independent of the means μ_g and μ_x.

(b) Repeat question (a) for the case of the causal affine Wiener estimator which satisfies the additional restriction that $h(k,j) = h(k-j) = 0$, $k < j$.

6.5 You have the measurement model

$$x_k = s_k^2 + w_k$$

where w_k is zero mean white noise of variance σ^2 and s_k is a w.s.s. zero mean Gaussian random sequence with acf $r_s(k) = a^k/(1-a^2)$, $k \geq 0$. w_k is independent of s_k.

(a) Find the quantities $E[s_i^2 s_{i-k}^2]$ and $E[s_i^2 s_{i-k}]$ (Hint: use the property that for any zero mean jointly Gaussian r.v.s W, X, Y, Z: $E[XYZ] = 0$ and $E[WXYZ] = E[WX]E[YZ] + E[WZ]E[XY] + E[WY]E[XZ]$)

(b) Using the results of (a) find the optimal affine non-causal Wiener estimator for s_k. Explain your result.

(c) Using the results of (a) find the optimal affine non-causal Wiener estimator for s_k^2.

6.6 Assume the measurement model

$$x_k = a s_k + v_k$$

where s_k and v_k are zero mean jointly w.s.s. and uncorrelated, and a is a random variable independent of s_k having mean μ_a and variance σ_a^2.

(a) Find the non-causal Weiner filter for estimating s_k.

(b) Find the MSE of the output of the non-causal Weiner filter. How does it behave as a function of μ_a and σ_a^2?

(c) Find the causal Weiner filter for estimating s_k. Specialize to the case where s_k is an AR(1) process as in Example 29 with pole at $-a$ and where v_k is white noise with variance σ^2.

6.7 For $|a| < 1$, use the geometric series formula $\sum_{k=0}^{\infty} c^k z^k = 1/(1-cz)$, $|cz| < 1$ to derive the following two results:

$$\left\{\frac{z^l}{1+az^{-1}}\right\}_+ = \begin{cases} (-a)^l \frac{1}{1+az^{-1}}, & l \geq 0 \\ \frac{z^l}{1+az^{-1}}, & l < 0 \end{cases} \tag{95}$$

and

$$\left\{\frac{z^l}{1+az}\right\}_+ = \begin{cases} 1, & l = 0 \\ 0, & l > 0 \end{cases} \tag{96}$$

Now apply these results to compute the Z-domain quantity (for $l \geq 0$)

$$\left\{\frac{z^l}{(1+az^{-1})(1+bz)}\right\}_+ \qquad |a|, |b| < 1$$

.

6.8 Let the measurement model be

$$x_k = h_k * s_k + v_k$$

where s_k and v_k are zero mean jointly w.s.s. and uncorrelated, and h_k is a causal and causally invertible filter with Z-transform $H(z)$.

(a) Find the non-causal Weiner filter for estimating s_k.

(b) Find the causal Weiner filter for estimating s_k.

(c) Compare the results for (a) and (b) to the estimator $h_k^{-1} * \hat{s}_k$ where $\hat{s}_k$ is the output of the standard Weiner filter for estimating s_k using the measurement model $x_k = s_k + v_k$ and h_k^{-1} is the inverse Z-transform of $1/H(z)$.

6.9 Let the measurement model be as in Example 29

$$x_k = s_k + v_k$$

where s_k and v_k are zero mean jointly w.s.s. and uncorrelated. It is desired to estimate $g_k = h_k * s_k$ where h_k is the causal FIR filter with transfer function

$$H(z) = 1 + \alpha z^{-1}, \quad \alpha \in (-1, 1)$$

. Assume that $\alpha \neq a \neq b$.

(a) Find the non-causal Weiner filter for estimating g_k.

(b) Find the causal Weiner filter for estimating g_k.

(c) Compare the results for (a) and (b) to the estimator $\hat{g}_k = h_k * \hat{s}_k$ where $\hat{s}_k$ is alternatively the output of the standard non-causal and causal Weiner filters, respectively, for estimating s_k.

6.10 The process s_k is a zero mean AR(2) process following the recursion

$$s_k = 0.8 s_{k-1} - 0.15 s_{k-2} + w_k$$

where w_k is zero mean white noise of variance 1.5 uncorrelated with $s_{k-1}, s_{k-2}, \ldots$. The observation is

$$x_k = s_k + v_k$$

where v_k is zero mean white noise with variance 0.5 independent of s_k.

(a) Express the AR(2) recursion in Z-transform domain as $\mathcal{Z}\{s_k\} = H(z)\mathcal{Z}\{w_k\}$ and use the input/output PSD relation $\mathcal{P}_s(z) = H(z)H(z^{-1})\mathcal{P}_w(z)$ to determine the PSD $\mathcal{P}_s(z)$ of s_k.

(b) Find the non-causal Wiener filter for estimating s_k.

(c) Find the causal Wiener filter for estimating s_k.

6.11 TBD

6.12 A common multipath model for a communications receiver is that the direct path signal plus an attenuated and delayed indirect path version of the signal are received in additive white noise:

$$x_k = s_k + b s_{k-1} + w_k$$

The objective is to estimate the signal s_k given a set of measurements $\{x_k\}_k$. In the following assume that w_k is zero mean white with variance σ_w^2, s_k is zero mean white with variance σ_s^2, b is a constant $|b| < 1$, and s_k, w_k are uncorrelated. You can assume that $(\sigma_s^2(1+b^2) + \sigma_w^2)/(\sigma_s^2 b) = 5/2$ if that helps simplify your answers to the following.

(a) Find the power spectral density (PSD) $\mathcal{P}_x(e^{j\omega})$ of x_k, the cross PSD $\mathcal{P}_{sx}(e^{j\omega})$ of s_k and x_k, and the spectral factorization of $\mathcal{P}_x(z)$, $z \in \mathbb{C}$.

(b) Find the optimal non-causal Wiener filter for estimating s_k.

(c) Find the optimal causal Wiener filter for estimating s_k.

6.13 In the derivation of the discrete time Kalman filter we assumed that the state noise $\underline{w}_k$ was uncorrelated with the measurement noise $v(k)$. In this problem we generalize the Kalman filter to the case where $E[\underline{w}_k v(l)] = V_{wv}\delta_{kl}$ where δ_{kl} is the kronecker delta function. Derive the Kalman filter equations.

6.14 The measurement equation is given by

$$x_k = s_k + v_k$$

where s_k satisfies the dynamic model ($|a| < 1$)

$$s_{k+1} = as_k + w_k, \quad s_0 = s_o$$

and v_k, w_k, s_o are uncorrelated, v_k and w_k are zero mean white noises with variances σ_v^2 and σ_w^2, respectively.

(a) Derive the Kalman filter equations.

(b) Derive the steady state state error covariance (variance) $R_{\tilde{s}}(\infty) \stackrel{\text{def}}{=} \lim_{k\to\infty} R_{\tilde{s}}(K|K-1)$ by setting $R_{\tilde{s}}(k+1|k) = R_{\tilde{s}}(k|k-1) = R_{\tilde{s}}(\infty)$ in the Kalman error covariance update formula and solving explicitly for $R_{\tilde{s}}(\infty)$. Find the corresponding steady state Kalman gain.

(c) By taking the Z-transform of the steady state Kalman state recursion show that the Kalman predictor $\hat{s}_{k+1|k}$ is the output of a LTI with input x_k.

(d) Compare the steady state Kalman predictor previously derived to the causal Wiener predictor based on the infinite past.

6.15 Let the random sequence $\{x_k\}$ be zero mean and wide sense stationary of the form

$$x_k = s_k + v_k$$

where s_k is a signal with PSD $\mathcal{P}_s$ and v_k is a white noise. You only get to measure the value of x_k for odd indices $k = 2n-1$, $n = -\infty, \ldots, \infty$. The objective is to estimate s_k at both odd and even time instants k. Note that when $v_k = 0$ the problem reduces to "filling in" the missing (even) data points.

(a) What is the system of Wiener-Hopf equations which must be satisfied by the optimal linear filter for estimating $\{s_k\}$ from the measurements? Is the solution to this system of equations time-invariant? If so find an expression for the optimal non-causal Wiener filter transfer function $H(z)$.

(b) Now assume that $s_k = as_{k-1} + w_{k-1}$ where $|a| < 1$ and w_k is white noise independent of v_k. Derive Kalman filter equations for recursively generating estimates $\hat{s}_{2n-1|2n-1}$ and $\hat{s}_{2n|2n-1}$ from the past measurements $\{x_{2k-1}\}_{k=1}^n$. Does the KF reduce to a linear time invariant filter as $n \to \infty$.

6.16 Derive the minimum mean square error expression (69) for the Wiener filter and use it to find the MSE of the optimal predictors of Examples 28 and 29.

6.17 TBD

6.18 In this exercise you will explore the extended Kalman filter (EKF) for non-linear state dynamics by extending the Bayesian derivation of the Kalman filter in Section 6.7.3. Similarly to that section we assume that the observations y_k obey a dynamical model, except that here we assume that the state can evolve *non-linearly*

$$\begin{aligned} y_k &= s_k + v_k \\ s_k &= \underline{c}^T \underline{\xi}_k \qquad\qquad , k = 0, 1, \ldots \\ \underline{\xi}_{k+1} &= \underline{g}(\underline{\xi}_k) + \mathbf{B}_k \underline{w}_k \end{aligned}$$

where $\underline{g}$ is a possibly non-linear p-dimensional function of the p-dimensional state vector $\underline{\xi}_k$, v_k and $\underline{w}_k$ are mutually independent zero mean temporally uncorrelated (white) noises which are *Gaussian* distributed with variance σ_v^2 and covariance matrix $\mathbf{R}_w$, respectively. All other assumptions on the model are identical to those made in Section 6.7.3. Let the posterior density of the state $\underline{\xi}_k$ given the observation sequence $\mathcal{Y}_l = \{y_l, y_{l-1}, \ldots, y_0\}$ up to time l be denoted $f_{\underline{\xi}_k|\mathcal{Y}_l}$.

In this exercise you will apply *Laplace's approximation* to the posterior distribution to obtain approximate state and covariance update recursions from Eq. (89). Laplace's approximation [17] asserts that the posterior is approximately Gaussian for large sample sizes

$$f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_k|\mathcal{Y}_k) \approx \frac{|\mathbf{F}_k|^{\frac{1}{2}}}{(2\pi)^{p/2}} \exp\left(-\frac{1}{2}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})^T \mathbf{F}_k (\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})\right)$$

where $\hat{\underline{\xi}}_{k|k} = \text{argmax}_{\underline{\xi}_k} f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_k|\mathcal{Y}_k)$ is the MAP estimator of $\underline{\xi}_k$ given past observations $\mathcal{Y}_k$ and $\mathbf{F}_k = \mathbf{F}(\hat{\underline{\xi}}_{k|k})$ is the $p \times p$ *observed Fisher information matrix (FIM)* where, for $\underline{u} \in \mathbb{R}^p$

$$\mathbf{F}(\underline{u}) = -\nabla_{\underline{u}}^2 \ln f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{u}|\mathcal{Y}_k).$$

(a) Using Laplace's approximation, and the approximation $\underline{g}(\underline{\xi}_k) = \underline{g}(\hat{\underline{\xi}}_{k|k}) + \nabla \underline{g}(\hat{\underline{\xi}}_{k|k})(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})$, in the integrand of the right hand side of (89), evaluate the integral by completion of the square.

(b) Using the results of part (a), and an analogous differentiation method to the one we used in the Bayesian derivation of the Kalman filter, generate a recursion $\hat{\underline{\xi}}_{k|k} \to \hat{\underline{\xi}}_{k+1|k+1}$ for the MAP state estimator and a recursion $\mathbf{F}_k \to \mathbf{F}_{k+1}$ for the observed FIM. These recursions represent the EKF filter. Represent your state estimator recursion in a form reminiscent of the Kalman filter, i.e.,

$$\hat{\underline{\xi}}_{k+1|k+1} = \underline{g}(\hat{\underline{\xi}}_{k|k}) + \Gamma_k \eta_k$$

where η_k is an analog to the Kalman innovation sequence and Γ_k is an analog to the Kalman Gain matrix (but which depends on $\hat{\underline{\xi}}_{k|k}$).

(c) Evaluate the EKF specified by the recursions found in (b) for the case of a scalar ($p = 1$) state ξ_k, scalar state noise w_k, scalar c, and the quadratic plant

$$g(\xi_k) = a\xi_k^2,$$

where $|a| < 1$. If the Fisher recursion is initialized by $F_{-1} > 0$ will the observed Fisher information remain F_k positive for all $k \geq 0$?

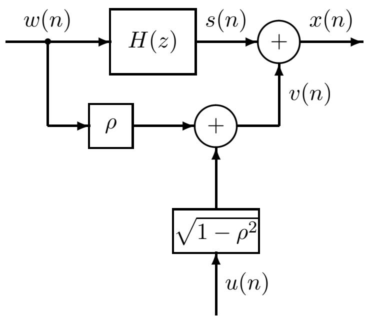
6.19 This is a continuation of the previous exercise. An approach to the EKF which does not require making the supplemental approximation $\underline{g}(\underline{\xi}_k) = \underline{g}(\hat{\underline{\xi}}_{k|k}) + \nabla \underline{g}(\hat{\underline{\xi}}_{k|k})(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})$ is to apply the Laplace approximation to the posterior of the successive pair of states

$$f_{\underline{\xi}_{k+1},\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_{k+1},\underline{\xi}_k|\mathcal{Y}_k) \approx \frac{|\mathbf{Q}_k|^{\frac{1}{2}}}{(2\pi)^p} \exp\left(-\frac{1}{2}\begin{bmatrix} \underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1} \\ \underline{\xi}_k - \hat{\underline{\xi}}_{k|k} \end{bmatrix}^T \mathbf{Q}_k \begin{bmatrix} \underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1} \\ \underline{\xi}_k - \hat{\underline{\xi}}_{k|k} \end{bmatrix}\right)$$

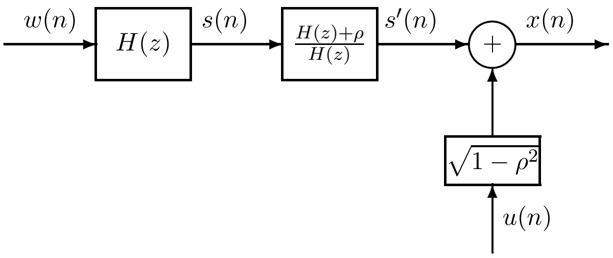
where $\mathbf{Q}_k = \mathbf{Q}(\hat{\underline{\xi}}_{k+1|k+1}, \hat{\underline{\xi}}_{k|k})$ is the $2p \times 2p$ observed FIM where, for $\underline{u}, \underline{v} \in \mathbb{R}^p$

$$\mathbf{Q}(\underline{u}, \underline{v}) = -\begin{bmatrix} \nabla^2_{\underline{u}} & \nabla_{\underline{u}}[\nabla_{\underline{v}}]^T \\ \nabla_{\underline{v}}[\nabla_{\underline{u}}]^T & \nabla^2_{\underline{v}} \end{bmatrix} \ln f_{\underline{\xi}_{k+1},\underline{\xi}_k|\mathcal{Y}_k}(\underline{u}, \underline{v}|\mathcal{Y}_k).$$

Find a set of EKF equations by using this approximation and compare to your solution in part (b) of Ex. 6.18. (You can assume that the state is 1 dimensional if you like).



(a) Block diagram for question 6.20



(b) Equivalent diagram for question 3

Figure 60: Block diagrams for question 6.20.

6.20 *Wiener filtering of a signal corrupted by noise which is correlated with the signal (R. Raich):* Consider the system in Fig. 60(a). The observations $x(n)$ are given by

$$x(n) = s(n) + v(n)$$

where $s(n)$ is an AR(1) random process ($H(z) = \frac{1}{1+az^{-1}}$, $|a| < 1$) given by

$$s(n) = (-a)s(n-1) + w(n)$$

and $v(n)$ is a noise which is partially correlated with the signal and is given by

$$v(n) = \rho w(n) + \sqrt{1-\rho^2}u(n),$$

where $0 \leq \rho \leq 1$. Both $u(n)$ and $w(n)$ are uncorrelated, zero mean, white noise processes with variances σ_u^2 and σ_w^2, respectively. To simplify the problem, an equivalent block diagram is presented in Fig. 60(b). (Hint $H(z) + \rho = \frac{1}{1+az^{-1}} + \rho$ can be simplified as $(1+\rho)\frac{1+bz^{-1}}{1+az^{-1}}$, where $b = \frac{\rho}{\rho+1}a$).

[(a)]

(a) *Non-causal Wiener Filtering:* Find the non-causal Wiener filter for $s(n)$ given $x(n)$. Express the filter in terms of ρ, $H(z)$, σ_w^2, and σ_u^2. (There is no need to substitute $H(z) = \frac{1}{1+az^{-1}}$ yet.)

(b) Explain and interpret what happens when $\rho = 0$ and $\rho = 1$. Obtain closed-form expressions for the Wiener filter in terms of σ_w^2, σ_u^2, a, and z (here, substitute $H(z) = \frac{1}{1+az^{-1}}$).

(c) *Causal Wiener Filtering:* Consider the case where $\rho = 1$. Find the whitening filter for the causal Wiener filter of $s(n)$ given $x(n)$.

(d) *Causal Wiener Filtering:* Consider the case where $\rho = 1$. Find the causal Wiener filter of $s(n)$ given $x(n)$.

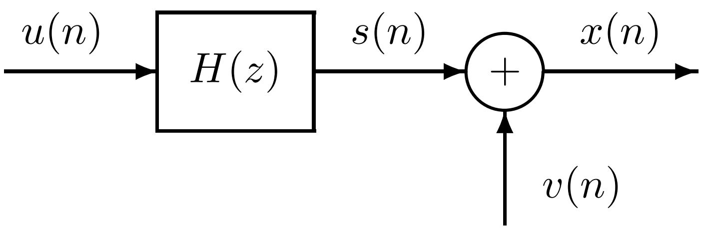


Figure 61: Block diagram for question 6.21

6.21 *Wiener/Kalman filtering of a moving average (MA) system (R. Raich):* Consider the system in Fig. 61. The observations $x(n)$ are given by

$$x(n) = s(n) + v(n)$$

where $s(n)$ is an MA(1) random process ($H(z) = 1 + az^{-1}$, $|a| < 1$) given by

$$s(n) = au(n-1) + u(n).$$

Both $v(n)$ and $u(n)$ are uncorrelated, zero mean, white noise processes with variances σ_v^2 and σ_u^2, respectively.

[(a)]

(a) *Non-causal Wiener Filtering:* Find the non-causal Wiener filter for $s(n)$ given $x(n)$.

(b) *Causal Wiener Filtering:*
- Find the whitening filter for the causal Wiener filter of $s(n)$ given $x(n)$. (Hint: find a similar relationship to the one in Exercise 6.3, which is applicable to this question.)
- Find the causal Wiener filter of $s(n)$ given $x(n)$.

(c) *Kalman Filtering:* Derive the Kalman filter for $s(n)$ given $x(n)$. (Hint: choose the state vector as $\underline{\xi}_n = [u(n), u(n-1)]^T$).

(d) *Kalman Filtering:* Find the steady state Kalman gain in terms of a, σ_v^2, and σ_u^2.

6.22 Let s_k be a signal satisfying the AR(1) recursion $s_{k+1} = as_k + w_k$, $-\infty < k < \infty$, where $a = 0.75$ and w_k is white noise with variance $\sigma_w^2 = 1$. The signal s_k is observed in uncorrelated white noise v_k of variance $\sigma_v^2 = 1$:

$$x_k = s_k + v_k, \quad -\infty \le k < \infty.$$

It is desired to estimate $g_k = s_k - \frac{1}{2}s_{k-1}$.

(a) Find the non-causal Wiener filter for estimating g_k.

(b) Find the causal Wiener filter for estimating g_k.

(c) If H_{ncw} is the non-causal Wiener filter for estimating s_k and $\hat{s}_k^{nc}$ is the output of this filter, compare the estimator of part (a) to the estimator $\hat{g}_k^{nc} = \hat{s}_k^{nc} - \frac{1}{2}\hat{s}_{k-1}^{nc}$. Comment on the difference between these estimators. Which one gives lower MSE?.

(d) If H_{cw} is the causal Wiener filter for estimating s_k and $\hat{s}_k^{c}$ is the output of this filter, compare the estimator of part (b) to the estimator $\hat{g}_k^{c} = \hat{s}_k^{c} - \frac{1}{2}\hat{s}_{k-1}^{c}$. Comment on the difference between these estimators. Which one gives lower MSE?

6.23 Available for observation is a zero mean w.s.s. signal $s(n)$ in additive white noise $v(n)$ of variance σ_v^2:

$$x(n) = s(n) + v(n),$$

where $s(n)$ and $v(n)$ are uncorrelated and $s(n)$ obeys the recursion

$$s(n) = as(n-2) + w(n)$$

and $|a| < 1$, $w(n)$ is zero mean white noise with variance 1. The objective is to estimate the signal $g(k) = s(k+\alpha)$ for $\alpha \geq 0$ a non-negative integer.

(a) Derive the power spectral density $\mathcal{P}_x(z)$of $x(n)$ in terms of a and σ_v^2.

(b) Plot the pole zero constellation of $\mathcal{P}_x$ and find the spectral factors $\mathcal{P}_x^+(z)$ and $\mathcal{P}_x^-(z)$.

(c) Derive the cross spectral density $\mathcal{P}_{gx}(z)$.

(d) Find the non-causal Wiener filter. To what does your filter reduce when $\sigma_v^2 = 0$ and $\alpha = 1$?

(e) Find the causal Wiener filter for estimating g_k when $\sigma_v^2 = 0$. Specialize to the case of $\alpha = 1$ and compare to your answer for part (d).

(f) Find the causal Wiener filter for $\sigma_v^2 > 0$.

6.24 Oftentimes there are outliers or missing observations that prevent direct implementation of the Kalman filter. In this problem you will explore one way to handle this scenario. You are given an observation model similar to the model (65) in Chapter 6.

$$\begin{aligned} x_k &= s_k + v_k \\ s_k &= \underline{c}^T \underline{\xi}_k \\ \underline{\xi}_{k+1} &= \mathbf{A}\underline{\xi}_k + \mathbf{B}\underline{w}_k \end{aligned}$$

where everything is the same as for (65) except that the observation noise v_k has variance $\sigma_v^2(k)$ that is itself a random variable with two states, a good state (small variance) and a bad state (large variance). This can be modeled by introducing a random switching variable b_k into the definition of $\sigma_v^2(k)$

$$\sigma^2(k) = b_k\sigma_0^2 + (1-b_k)\sigma_1^2$$

where $\sigma_0^2 < \sigma_1^2$, and $\{b_k\}$ are i.i.d. Bernoulli random variables with $P(b_k = 1) = p$ and $P(b_k = 0) = 1-p$. We assume that the b_k's are independent of the states $\{\underline{\xi}_k\}$. This model introduces missing observations by taking the limit of expressions for the optimal predictors as $\sigma_1^2 \to \infty$. To handle the case of unobserved switching variables we will assume in (b), (c) and (d) that the noises v_k, $\underline{w}_k$ and the initial state $\underline{\xi}_0$ are independent Gaussian distributed.

(a) For the case that the measurements x_k and the switching variables b_k are both observed, i.e. the outlier indices are known, give the filtering equations for the Kalman filter estimator of s_k given $(x_k, b_k), (x_{k-1}, b_{k-1}), \ldots, (x_0, b_0)$. Be sure to specify the state estimator updates, the Kalman gain, and the error covariance updates in terms of b_k.

(b) To what do your equations in part (a) converge as $\sigma_1^2 \to \infty$? How about as $\sigma_1^2 \to \sigma_0^2$? How do your state update equations compare to those of the standard Kalman filter?

(c) Show that the conditional density $f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}|\underline{\xi}_k)$ has the form of a multivariate Gaussian density and specify the mean and covariance matrix of this conditional density.

(d) For the case of unobserved switching variables, find an expression for the instantaneous state likelihood function $f_{x_k|\underline{\xi}_k}(x|\underline{\xi})$ in terms of the switching probability p. (Hint: you will need to "integrate out" the switching variables).

(e) Use Bayes theorem to show that the posterior density of the state, $\rho_k(\underline{\xi}_k) \stackrel{\text{def}}{=} f_{\underline{\xi}_k|\mathcal{X}_k}(\underline{\xi}_k|\mathcal{X}_k)$, where $\mathcal{X}_k = \{x_k, x_{k-1}, \ldots, x_0\}$, obeys the recursion

$$\rho_{k+1}(\underline{\xi}_{k+1}) \quad = \quad \frac{f_{x_{k+1}|\underline{\xi}_{k+1}}(x_{k+1}|\underline{\xi}_{k+1})}{f_{x_{k+1}|\mathcal{X}_k}(x_{k+1}|\mathcal{X}_k)} \int f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k)\rho_k(\underline{\xi}_k)d\underline{\xi}_k.$$

where the density $f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k)$ has the form of a multivariate Gaussian density. This recursion can be used to generate an update equation for the conditional mean state estimator (see part (f)).

(f) Use the recursion (e) to find a time update of the conditional mean $\hat{s}_{k|k} = E[s_k|\mathcal{X}_k]$ of the form $\hat{s}_{k|k} \to \hat{s}_{k+1|k+1}$. You may specialize to the AR(1) model for $s_k = \underline{\xi}_k$ as in homework problem (6.14) and assume low SNR, e.g., $a^2/\sigma_1^2, \sigma_w^2/\sigma_1^2 << 1$.

End of chapter

7 FUNDAMENTALS OF DETECTION

Next we treat the problem of detection: this is equivalent to estimation when there are only a small number of possible values for the unknown parameter $\underline{\theta}$. Some may argue that detection has a simpler and more elegant theory than estimation, and this might be true depending on the "eye of the beholder." However, as we will see in the next chapter detection and estimation are closely linked when there exist unknown nuisance parameters that can confound our detection algorithms.

We will cover the following in this chapter

* Optimal detection theory

* Bayesian approach to detection

* Frequentist approach to detection

* Reciever Operating Characteristic (ROC) curves

* Multiple hypothesis testing

Example 31 *A motivating radar example*

We start with a practical example to motivate the detection theory to come later. Assume that you make a continuous time measurement $x(t)$ over a time interval $[0, T]$ and you wish to decide whether $x(t)$ is noise alone

$$x(t) = w(t), \quad 0 \leq t \leq T$$

or whether it is signal plus noise

$$x(t) = \theta s(t - \tau) + w(t), \quad 0 \leq t \leq T.$$

Here we assume

* $s(t)$ is a known signal that may or may not be present

* $w(t)$ is a zero mean Gaussian white noise with known power spectral density level $N_o/2$

*τ is a known time delay, $0 \leq \tau \ll T$

* $\int_0^T |s(t-\tau)|^2 dt = \int_0^T |s(t)|^2 dt$ is the signal energy

* $\theta \in \{0, 1\}$ unknown nuisance parameter

The detection objective is to decide whether the signal is present or not, and to do this with minimum average number of decision errors.

There is a common notation that has been developed for stating the detection hypotheses: "no signal present" (H_0) vs "signal present" (H_1)

$$\begin{array}{lcl} H_0 : x(t) = w(t) & & H_0 : \theta = 0 \\ & \Leftrightarrow & \\ H_1 : x(t) = s(t-\tau) + w(t) & & H_1 : \theta = 1 \end{array}$$

Without trying to choose a decision function to optimize any particular detection performance criterion - we will do this later - two methods of detection could be considered, shown in Fig. 62:

Energy-threshold detector:

$$y = \int_0^T |x(t)|^2 dt \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

Filter-threshold detector

$$y = \int_0^T h(T-t)x(t)dt \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

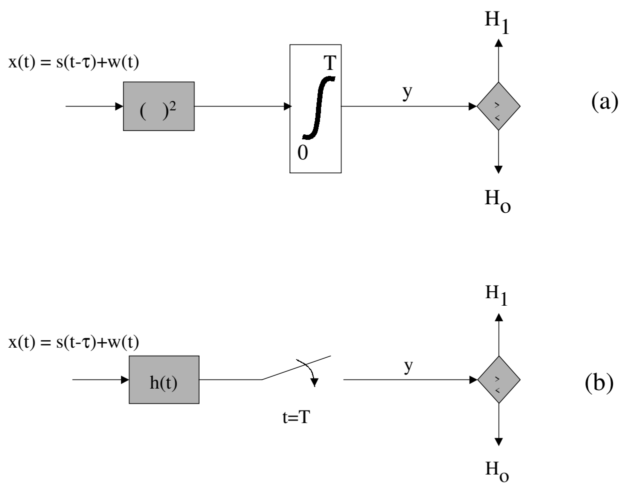


Figure 62: *(a) energy-threshold detector, (b) filter-threshold detector*

ERROR ANALYSIS:

Referring to Fig. 63, there are two types of decision error to be concerned about:

FALSE ALARM: $y > \eta$ when no signal present

MISS: $y < \eta$ when signal present

We can easily compute the conditional probabilities of these errors when there is no signal present and when there is a signal present, respectively:

$$\begin{aligned} P_F &= P(\text{say signal}|\text{no signal}) = \int_{y>\eta} f(y|\text{no signal})dy \\ P_M &= P(\text{say no signal}|\text{signal}) = \int_{y\leq\eta} f(y|\text{signal})dy \end{aligned}$$

TWO QUESTIONS

Q1: Is there an optimal way of trading off P_M for P_F?

Q2: Can we optimize the filter $h(\cdot)$ in the filter-threshold detector to give optimal tradeoff?

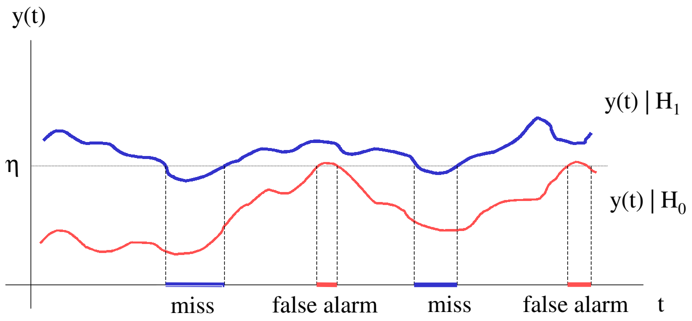


Figure 63: *Repeated tests of radar produce sequence of y_i's. One sequence has signal present (H_1) and one has signal absent (H_0).*

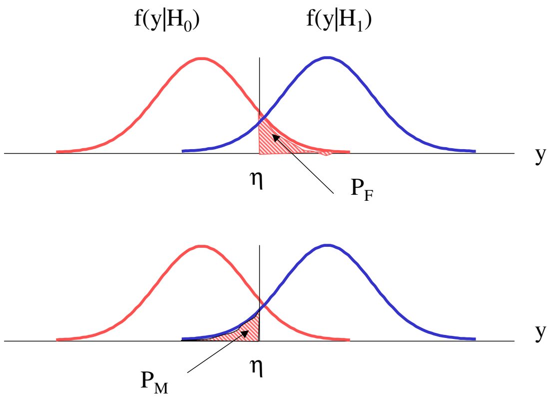


Figure 64: *Miss probability P_M and false alarm probability P_F for the radar example. Note that decreasing one type of error by changing the decision threshold η is necessarily accompanied by increasing the other type of error.*

We will defer the answer to Q1 till later. The filter in the filter-threshold detector can be chosen to optimize a design criterion. As a large overlap of the two densities in Fig. 64 makes the tradeoff much worse between the two types of error, a reasonable strategy would be to choose the filter $h(\cdot)$ to minimize this overlap. A measure of the amount of overlap is the *deflection*

$$d^2 = \frac{|E[y|\text{signal}] - E[y|\text{no signal}]|^2}{\text{var}(y|\text{no signal})}.$$

Large values of d^2 translate into well separated densities $f(y|H_0)$ and $f(y|H_1)$ with low overlap. Our objective should be to maximize d^2 and thus minimize the overlap.

We can easily compute the deflection for our radar example. Note that the presence of the signal produces shift in mean but not in the variance

$$\begin{aligned}
E[y|\text{no signal}] &= 0 \\
E[y|\text{signal}] &= \int_0^T h(T-t)s(t-\tau)dt \\
\text{var}[y|\text{no signal}] &= N_o/2 \int_0^T |h(t)|^2 dt.
\end{aligned}$$

Then, applying the Cauchy-Schwarz inequality

$$d^2 = \frac{2}{N_o} \frac{\left|\int_0^T h(T-t)s(t-\tau)dt\right|^2}{\int_0^T |h(T-t)|^2 dt} \leq \frac{2}{N_o} \underbrace{\int_0^T |s(t-\tau)|^2 \, dt}_{\int_0^T |s(t)|^2 = \|s\|^2}$$

with "=" if and only if $h(T-t) = as(t-\tau)$ for some constant a.

$\Rightarrow$ obtain "matched filter" solution as optimal deflection filter:

$$h(t) = s(T + \tau - t)$$

CASE of $s(\tau)$ = a short duration "pulse"

* $\int_0^T |s(t-\tau)|^2 dt$ does not depend on τ

* optimal detector can be implemented as:

$$\begin{aligned}
y &= \int_0^T s(t-\tau)x(t)dt \\
&= \int_{-\infty}^{\infty} s(t-\tau)x(t)dt \\
&= s(-t) * x(t)\big|_{t=\tau}
\end{aligned}$$

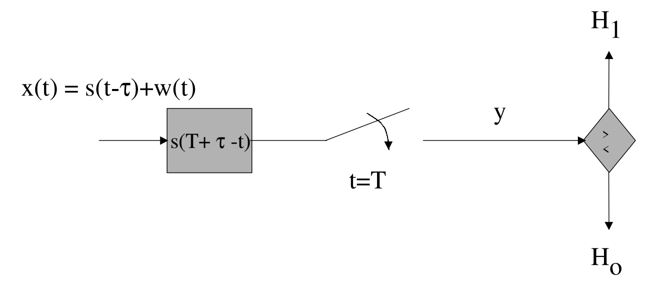


Figure 65: *SNR optimal receiver implemented as a matched filter reciever for delayed signal in noise.*

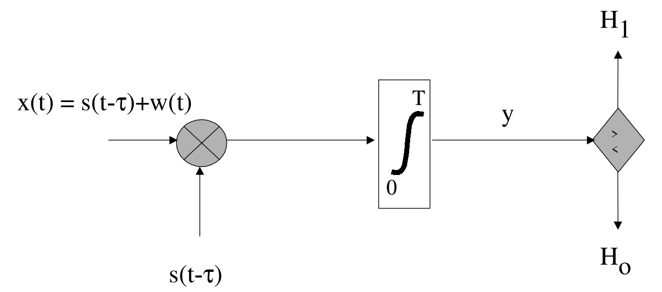


Figure 66: *SNR optimal receiver implemented as a correlator reciever for delayed signal in noise.*

7.1 THE GENERAL DETECTION PROBLEM

Let's now turn to the general detection problem. We have the following setup, as before:

X a measured random variable, random vector, or random process

$x \in \mathcal{X}$ is a realization of X

$\theta \in \Theta$ are unknown parameters

$f(x; \theta)$ is p.d.f. of X (a known function)

Two distinct hypotheses on θ

$$\theta \in \Theta_0, \quad or \quad \theta \in \Theta_1$$

Θ_0, Θ_1 is partition of Θ into two disjoint regions

$$\Theta_0 \cup \Theta_1 = \Theta, \quad \Theta_0 \cap \Theta_1 = \{\text{empty}\}$$

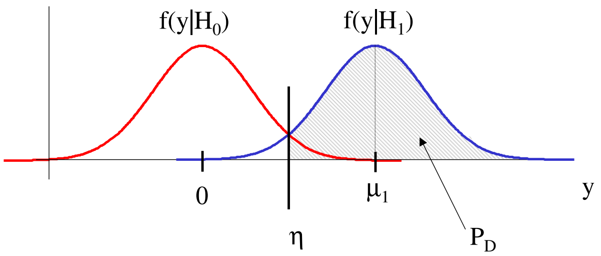


Figure 67: *Detection probability $P_D = 1 - P_M$ for the radar example.*

NOTATION:

$$\begin{array}{ccc} H_0 : \theta \in \Theta_0 & & H_0 : X \sim f(x;\theta), \;\; \theta \in \Theta_0 \\ & \Leftrightarrow & \\ H_1 : \theta \in \Theta_1 & & H_1 : X \sim f(x;\theta), \;\; \theta \in \Theta_1 \end{array}$$

H_0: the null hypothesis, noise alone hypothesis

H_1: the alternative hypothesis, signal present hypothesis

As the true hypothesis is not under our control it is often called the "true state of nature."

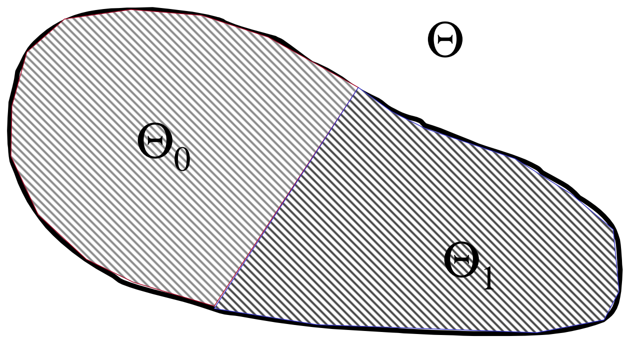


Figure 68: *The detector must decide on the region of* Θ *that contains the unknown parameter* θ*.*

7.1.1 SIMPLE VS COMPOSITE HYPOTHESES

When θ can take on only two values and Θ_0 and Θ_1 are singleton sets, the hypotheses are said to be *simple*.

$\Theta = \{\theta_0, \theta_1\}, \quad \Theta_0 = \{\theta_0\}, \quad \Theta_1 = \{\theta_1\}$.

In this case the p.d.f. $f(x;\theta)$ is completely known given either H_0 or H_1

If the hypotheses are not simple then at least one of Θ_1 or Θ_0 is not a singleton and is said to be composite. Simple hypotheses are much easier to deal with and one is lucky to encounter them in practical problems!

7.1.2 THE DECISION FUNCTION

Detection objective: design a decision rule (test function)

$$\phi(x) = \begin{cases} 1, & \text{decide } H_1 \\ 0, & \text{decide } H_0 \end{cases}. \tag{97}$$

The test function $\phi(x)$ maps $\mathcal{X}$ to the decision space $\{0,1\}$ for deciding H_0 and H_1. The function $\phi(x)$ induces a partition of $\mathcal{X}$ into decision regions

$$\mathcal{X}_0 = \{x : \phi(x) = 0\}, \quad \mathcal{X}_1 = \{x : \phi(x) = 1\}$$

FALSE ALARM AND MISS ERRORS

False alarm and miss probabilities associated with the test function ϕ can be expressed simply:

$$P_F(\theta) = E_\theta[\phi] = \int_{\mathcal{X}} \phi(x) f(x;\theta) dx, \quad \theta \in \Theta_0,$$

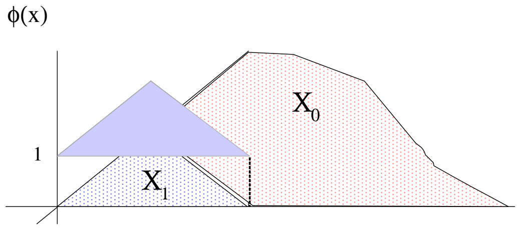


Figure 69: *Test function separates measurement space into two decision regions $\mathcal{X}_0$ and $\mathcal{X}_1$ (the region under the raised platform).*

$$P_M(\theta) = E_\theta[1-\phi] = \int_{\mathcal{X}} [1-\phi(x)] f(x;\theta) dx, \quad \theta \in \Theta_1,$$

where in the expectation expressions the reader must interpret $\phi = \phi(X)$ as a random variable. Equivalently

$$\begin{aligned} P_F(\theta) &= \int_{\mathcal{X}_1} f(x|\theta) dx, \quad \theta \in \Theta_0 \\ P_M(\theta) &= 1 - \int_{\mathcal{X}_1} f(x|\theta) dx, \quad \theta \in \Theta_1 \end{aligned}$$

The probability of correctly deciding H_1 is called the (correct-) detection probability:

$$1 - P_M(\theta) = P_D(\theta) = E_\theta[\phi], \quad \theta \in \Theta_1$$

We give separate treatment for case of random and non-random θ

7.2 BAYES APPROACH TO DETECTION

There are three elements involved in taking a Bayesian approach. One must:

1. Assign a prior $f(\theta)$ density for θ

2. Assign a cost or risk to wrong decisions

* c_{ij} = cost of deciding H_i when H_j is true

3. Find and implement decision rule which has minimum average risk

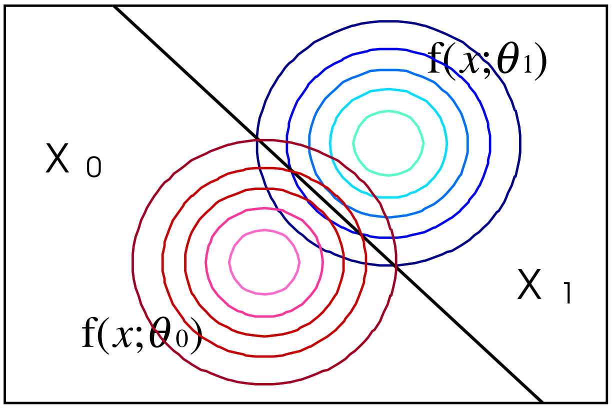


Figure 70: *Illustration of decision regions* χ_0 *and* χ_1 *for deciding* H_0 *and* H_1 *for an observation* x *in the plane. Also shown are constant contours of the* H_0 *and* H_1 *densities* $f(x;\theta_0)$ $f(x;\theta_1)$*. False alarm probability* P_F *is integral of* $f(x;\theta_0)$ *over* χ_1*, miss probability* P_M *is integral of* $f(x;\theta_1)$ *over* χ_0*, and detection probability* P_D *is integral of* $f(x;\theta_1)$ *over* χ_1*.*

7.2.1 ASSIGNING PRIOR PROBABILITIES

Obtain prior probabilities on H_0, H_1

$$\begin{aligned} P(H_0) &= P(\theta \in \Theta_0) = \int_{\Theta_0} f(\theta)d\theta \\ P(H_1) &= P(\theta \in \Theta_1) = \int_{\Theta_1} f(\theta)d\theta \end{aligned}$$

with $P(H_0) + P(H_1) = 1$

In this case we can compute conditional p.d.f.'s given H_0 and H_1 by integrating over θ

$$\begin{aligned} f(x|H_0) &= \frac{\int_{\Theta_0} f(x|\theta)f(\theta)d\theta}{P(H_0)} \\ f(x|H_1) &= \frac{\int_{\Theta_1} f(x|\theta)f(\theta)d\theta}{P(H_1)} \end{aligned}$$

7.2.2 MINIMIZATION OF AVERAGE RISK

We first define the cost or risk matrix:

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{10} \\ c_{01} & c_{00} \end{bmatrix}.$$

We will assume throughout that $c_{ii} \leq c_{ij}$, i.e. the cost of making a correct decision is less than that of making an incorrect one. The actual cost incurred for a given realization of X, which we will call C, is a function of the outcome $\phi(X)$ of the test and a function of the true state, H_0 or H_1, of nature. The cost $C \in \{c_{11}, c_{10}, c_{01}, c_{00}\}$ is therefore a random variable and we can seek decision rules that minimize its average value, called the "average risk" associated with the decision function.

We adopt the following "Bayes" design criterion: Select ϕ, equivalently $\mathcal{X}_0$ and $\mathcal{X}_1$, to minimize average risk, equal to the statistical expectation $E[C]$ of the incurred cost C

$$\begin{aligned} E[C] &= c_{11}P(\text{say } H_1|H_1)P(H_1) + c_{00}P(\text{say } H_0|H_0)P(H_0) \\ &\quad + c_{10}P(\text{say } H_1|H_0)P(H_0) + c_{01}P(\text{say } H_0|H_1)P(H_1) \end{aligned} \tag{98}$$

Define the Bayesian false alarm and miss probabilities

$$\begin{aligned} P_F &= \int_{\mathcal{X}_1} f(x|H_0)dx = P(\text{say } H_1|H_0) \\ P_M &= 1 - \int_{\mathcal{X}_1} f(x|H_1)dx = P(\text{say } H_0|H_1) \end{aligned} \tag{99}$$

These differ from the probabilities $P_F(\theta)$ and $P_M(\theta)$ defined above since they denote error probabilities that involve averages of θ over Θ_0 and Θ_1. With these definitions we can express (98) in equivalent form

$$\begin{aligned} E[C] &= c_{00}P(H_0) + c_{11}P(H_1) \\ &\quad + [c_{01} - c_{11}]P(H_1)P_M + [c_{10} - c_{00}]P(H_0)P_F \end{aligned}$$

Observe: $E[C]$ linear in $P_M, P_F, P(H_1), P(H_0)$ for any fixed decision rule ϕ. This will become important when we start comparing performances of different decision rules so take note!

7.2.3 OPTIMAL BAYES TEST MINIMIZES $E[C]$

Using the integral representation (99) allows us to rewrite $E[C]$ explicitly as function of decision region $\mathcal{X}_1$

$$\begin{aligned} &E[C] = c_{00}P(H_0) + c_{01}P(H_1) \\ &+ \int_{\mathcal{X}_1} \left([c_{10} - c_{00}]P(H_0)f(x|H_0) - [c_{01} - c_{11}]P(H_1)f(x|H_1)\right)\, dx \end{aligned}$$

The solution is now obvious: if we had a choice to assign a candidate point x to $\mathcal{X}_1$ or to $\mathcal{X}_0$ we would choose $\mathcal{X}_1$ only when it decreased the average risk, i.e., made the integrand negative. Thus, assign x to $\mathcal{X}_1$ if

$$[c_{10} - c_{00}]P(H_0)f(x|H_0) < [c_{01} - c_{11}]P(H_1)f(x|H_1)$$

and assign x to $\mathcal{X}_0$ otherwise. This obvious solution can be formally proved by using an *exchange argument*: assume that a point x for which the integrand was positive was assigned to $\mathcal{X}_1$ and reason that you could always decrease the integral by reassigning the point to $\mathcal{X}_0$.

When $c_{10} > c_{00}$ and $c_{01} > c_{11}$ the optimal test is therefore the Bayes likelihood ratio test (BLRT)

$$\Lambda_B(x) := \frac{f(x|H_1)}{f(x|H_0)} \overset{H_1}{\underset{H_0}{\gtrless}} \eta$$

where η is the optimal Bayes threshold

$$\eta = \frac{[c_{10} - c_{00}]P(H_0)}{[c_{01} - c_{11}]P(H_1)}$$

The random variable $\Lambda_B(X)$ is called the Bayes likelihood ratio test (BLRT) statistic. Note that the costs and the prior probability $p = P(H_0) = 1 - P(H_1)$ only influence the BLRT through the threshold η, the Bayes likelihood ratio statistic $\Lambda_B(x)$ does not depend on p.

7.2.4 MINIMUM PROBABILITY OF ERROR TEST

Consider the special case of $c_{00} = c_{11} = 0$ and $c_{01} = c_{10} = 1$. This turns the average risk into the prob. error criterion

$$E[C] = P_M P(H_1) + P_F P(H_0) = P_e$$

which is minimized by the LR test

$$\frac{f(x|H_1)}{f(x|H_0)} \overset{H_1}{\underset{H_0}{\gtrless}} \frac{P(H_0)}{P(H_1)}.$$

Using Bayes rule you can easily see that this is equivalent to the "Maximum a posteriori" (MAP) test

$$\frac{P(H_1|x)}{P(H_0|x)} \overset{H_1}{\underset{H_0}{\gtrless}} 1.$$

7.2.5 PERFORMANCE OF BAYES LIKELIHOOD RATIO TEST

To simplify notation we define $\overline{C} = E[C]$. Let the minimum of risk $\overline{C}$, attained by the BLRT, be denoted $\overline{C}^*$

$$\begin{aligned} \overline{C}^* &= c_{00}P(H_0) + c_{11}P(H_1) \\ &\quad + [c_{01} - c_{11}]P(H_1)P_M^*(\eta) + [c_{10} - c_{00}]P(H_0)P_F^*(\eta) \end{aligned}$$

where

$$P_F^*(\eta) = P(\Lambda_B > \eta|H_0), \qquad P_M^*(\eta) = P(\Lambda_B \leq \eta|H_1).$$

Viewing $\overline{C}^* = \overline{C}^*(p)$ as a function of p, the minimum risk describes a performance curve (Fig. 71) as a function of $p = P(H_0)$ that is called the *minimum risk curve.* Note that this curve does not specify the performance of any single test function as a function of p; recall that the average risk of any specified test is linear in p. Rather it specifies the risk that would be attainable if the different optimal BLRT's were implemented for different values of p, i.e. different BLRT thresholds. Thus the minimum risk curve prescribes a lower bound on the average risk attained by any test for any value of p.

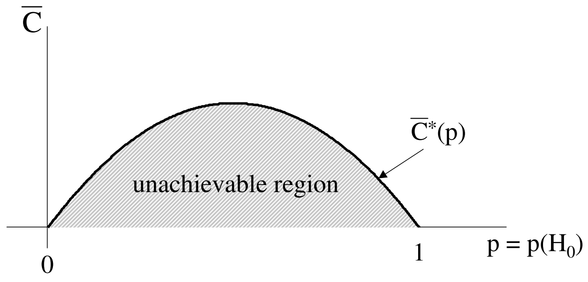


Figure 71: *The minimum risk curve associated with optimal BLRTs specifies an achievable lower bound on average risk of any test.*

7.2.6 MIN-MAX BAYES DETECTOR

In many cases the true value of p is unknown to the experimenter or designer of the test. Therefore, the optimal threshold of the BLRT cannot be implemented. As any specified test, even a BLRT with fixed threshold, has a linear average risk it might incur an unacceptably large average risk as p approaches either 0 or 1 (see straight line in Fig. 72). A sensible alternative in such a situation is for the designer to adopt a minimax strategy: if nature gets to select the true p then we should select a test to minimize worst case average risk

$$\overline{C}_{minimax} = \max_{p\in[0,1]} \overline{C}(p)$$

It is intuitively obvious from Fig. 72 that the minimax test must be an optimal Bayes test, i.e., a test whose average risk line is tangent to the minimum risk curve, implemented with a threshold η^* which makes $\overline{C}$ a horizontal line, i.e. the slope of $\overline{C}$ should be zero. Thus we have the following minimax optimality condition

$$\begin{aligned}\overline{C} &= \underbrace{[c_{00}(1-P_F^*(\eta)) + c_{10}P_F^*(\eta) - c_{11}(1-P_M^*(\eta)) - c_{01}P_M^*(\eta)]}_{=0}\, p \\ &\quad + c_{11}(1-P_M^*(\eta)) + c_{01}P_M^*(\eta)\end{aligned}$$

where $P_F^*(\eta)$ and $P_M^*(\eta)$ are the Bayesian false alarm and miss probabilities of the BLRT implemented with threshold η.

In the special case $\overline{C} = P_e$: $c_{00} = c_{11} = 0$, $c_{10} = c_{01} = 1$ we obtain the minimax condition on the MAP test:

$$\overline{C} = \underbrace{[P_F^*(\eta) - P_M^*(\eta)]}_{=0}\, p + P_M^*(\eta)$$

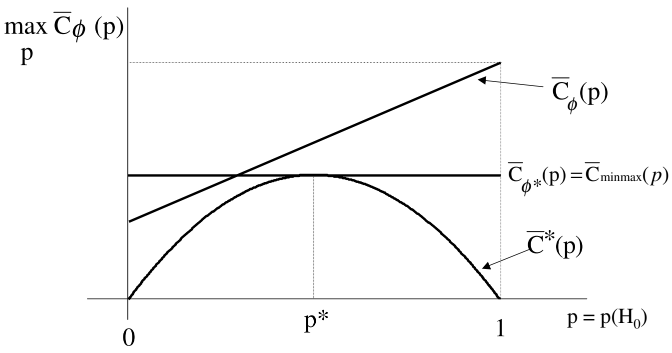


Figure 72: *Power curve of any fixed test ϕ is a straight line. The minimax optimal test ϕ^* has a horizontal power curve which is tangent to the minimum risk, denoted $\overline{C}^*(p)$, at its maximum.*

This implies that η should be selected so as to ensure the "equalization" condition is satsified

$$P_F^*(\eta) = P(\Lambda_B > \eta | H_0) = P(\Lambda_B \leq \eta | H_1) = P_M^*(\eta).$$

Denoting this minimax value of η as η^*, and noting that the designer can choose a threshold by choosing (guessing) a value of p, the minimax threshold is related to a minimax choice p^* through the relation $\eta^* = p^*/(1-p^*)$.

7.2.7 EXAMPLES

Example 32 *Radar example revisited*

Objective: Given the matched filter output y find the Bayes optimal detector.

1. Assume that $P(H_0) = P(H_1) = \frac{1}{2}$
2. Recall that y is

$$y = \int_0^T s(t)x(t)dt$$

which is a realization of a Gaussian random variable Y having means and variances

$$E[Y|H_0] = 0, \qquad \text{var}[Y|H_0] = N_o/2 \int_0^T |s(t)|^2 dt = \sigma_0^2$$

$$E[Y|H_1] = \int_0^T |s(t)|^2 dt = \mu_1, \quad \text{var}[Y|H_1] = N_o/2 \int_0^T |s(t)|^2 dt = \sigma_0^2$$

Bayes LR test is

$$\begin{aligned}
\Lambda_B(y) &= \frac{\frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2\sigma_0^2}(y-\mu_1)^2}}{\frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2\sigma_0^2}y^2}} \\
&= e^{y\mu_1/\sigma_0^2 - \frac{1}{2}\mu_1^2/\sigma_0^2} \overset{H_1}{\underset{H_0}{\gtrless}} \eta = 1
\end{aligned}$$

LR test statistic $\Lambda_B(Y)$ is a monotone function of Y since $\mu_1 > 0$.
Equivalent test is filter-threshold detector

$$Y \overset{H_1}{\underset{H_0}{\gtrless}} \gamma = \tfrac{1}{2}\mu_1$$

Performance of Bayes LRT for radar example

$$\begin{aligned}
P_F &= P(Y > \gamma | H_0) \\
&= P(\underbrace{Y/\sigma_0}_{\mathcal{N}(0,1)} > \gamma/\sigma_0 | H_0) \\
&= \int_{\gamma/\sigma_0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du \\
&= 1 - \mathcal{N}(\gamma/\sigma_0) := Q(\gamma/\sigma_0)
\end{aligned}$$

and

$$\begin{aligned}
P_M &= P(Y < \gamma | H_1) \\
&= P(\underbrace{(Y - \mu_1)/\sigma_0}_{\mathcal{N}(0,1)} > (\gamma - \mu_1)/\sigma_0 | H_1) \\
&= \mathcal{N}((\gamma - \mu_1)/\sigma_0)
\end{aligned}$$

Note: since the standard Gaussian p.d.f. is symmetric
$\mathcal{N}(-u) = 1 - \mathcal{N}(u)$ and thus, since $\gamma = \mu_1/2$

$$P_M = P_F = 1 - \mathcal{N}(\mu_1/(2\sigma_0)).$$

We thus conclude that the Bayes threshold γ is actually minimax! Therefore the probability of error reduces to

$$\begin{aligned}
P_e &= P_M \tfrac{1}{2} + P_F \tfrac{1}{2} \\
&= P_F
\end{aligned}$$

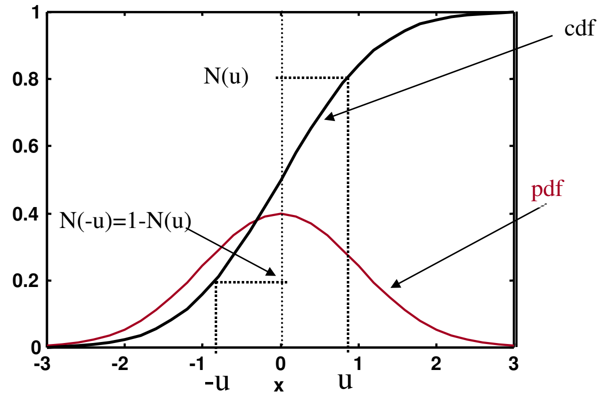


Figure 73: *CDF $\mathcal{N}(u)$ of symmetric Gaussian density*

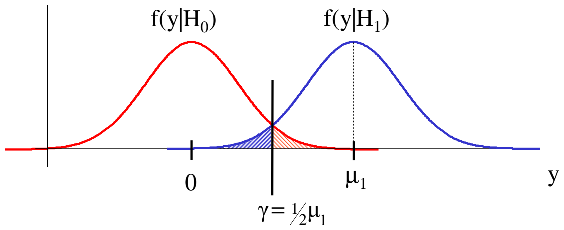


Figure 74: *Equally likely hypotheses have minmax threshold $\gamma = \frac{1}{2}\mu_1$ for problem of detection of shift in mean of a Gaussian r.v.*

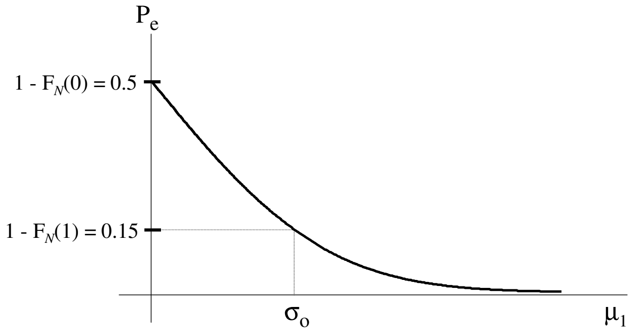


Figure 75: *Error probability curve of Bayes LRT as function of* $\mu_1 = \|s\|^2$.

7.3 TESTING MULTIPLE HYPOTHESES

We measure x having conditional p.d.f. $f(x|\theta)$

* $\theta \in \Theta$

Consider partition $\Theta_1, \ldots, \Theta_M$ of parameter space

OBJECTIVE: To test M hypotheses on θ

$$\begin{array}{rcl} H_1 : \theta & \in & \Theta_1 \\ \vdots & \vdots & \vdots \\ H_M : \theta & \in & \Theta_M \end{array}$$

DECISION FUNCTION:

$$\underline{\phi}(x) = [\phi_1(x), \ldots, \phi_M(x)]^T$$

where

$$\phi_i(x) \in \{0, 1\}, \qquad \sum_{i=1}^{M} \phi_i(x) = 1$$

NOTE: decision function specifies a partition of measurement space $\mathcal{X}$ into M decision regions

$$\mathcal{X}_i = \{x : \phi_i(x) = 1\}, \quad i = 1, \ldots, M$$

BAYES APPROACH: has three elements:

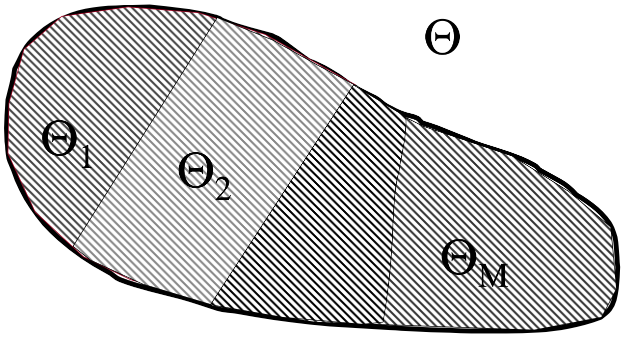


Figure 76: *Partition of* Θ *into* M *different regions*

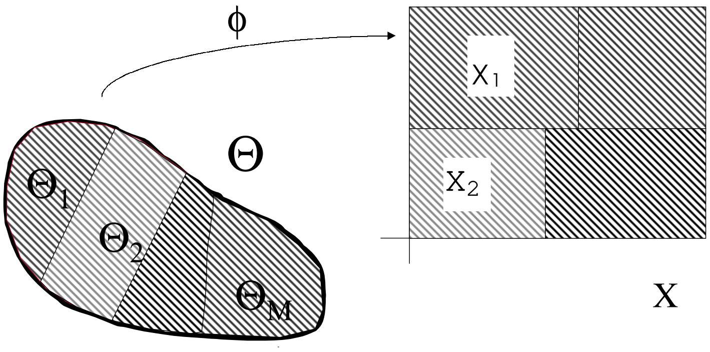


Figure 77: *Partition of parameter space into* M *hypotheses is equivalent (via the test function* $\underline{\phi}(x)$*) to partition of* $\mathcal{X}$ *into* M *regions.*

1. Assign a prior $f(\theta)$ density for θ
2. Assign costs to wrong decisions

c_{ij} = cost of deciding H_i when H_j is true

3. Find and implement decision rule which has minimum average cost

7.3.1 PRIOR PROBABILITIES

Obtain prior probabilities on H_i, $i = 1, \ldots, M$

$$P(H_i) = P(\theta \in \Theta_i) = \int_{\Theta_i} f(\theta) d\theta$$

with $\sum_{i=1}^M P(H_i) = 1$
We now have conditional p.d.f.s

$$f(x|H_i) = \frac{\int_{\Theta_i} f(x|\theta) f(\theta) d\theta}{P(H_i)}$$

$\Rightarrow$ Thus we have reduced the composite hypotheses above to the following simple hypotheses

$$\begin{array}{rcl} H_1 : X & \sim & f(x|H_1) \\ \vdots & \vdots & \vdots \\ H_M : X & \sim & f(x|H_M) \end{array}$$

where H_i has prior probability $P(H_i)$

7.3.2 MINIMIZE AVERAGE RISK

Cost or risk matrix is now $M \times M$:

$$\mathbf{C} = \begin{bmatrix} c_{11} & \cdots & c_{1M} \\ \vdots & \ddots & \vdots \\ c_{M1} & \cdots & c_{MM} \end{bmatrix}$$

Design Criterion: Select $\underline{\phi}$, equivalently $\{\mathcal{X}_i\}_{i=1}^M$, to minimize average risk $E[C] = \overline{C}$

$$\overline{C} = \sum_{i,j=1}^M c_{ij} P(\text{say } H_i | H_j) P(H_j)$$

Specialize to case

* $c_{ii} = 0$
* $c_{ij} = 1,\ i \neq j$

Then we have $\overline{C} = P_e$

$$\begin{aligned}
\overline{C} &= \sum_{i,j:i\neq j} P(\text{say } H_i|H_j)P(H_j) \\
&= 1 - \textstyle\sum_{i,j:i=j} P(\text{say } H_i|H_j)P(H_j) \\
&= 1 - \textstyle\sum_{i,j:i=j} P(X \in \mathcal{X}_i|H_i)P(H_i) \\
&= 1 - \sum_{i=1}^{M} \int_{\mathcal{X}_i} f(x|H_i)dx\ P(H_i)
\end{aligned}$$

Observe: to make $\overline{C}$ as small as possible

$$x \in \mathcal{X}_i \quad \Leftrightarrow \quad f(x|H_i)P(H_i) \geq f(x|H_j)P(H_j), \quad j \neq i$$

Or in terms of decision function:

$$\phi_i(x) = \begin{cases} 1, & f(x|H_i)P(H_i) \geq f(x|H_j)P(H_j) \\ 0, & o.w. \end{cases}$$

Shorthand notation

$$\hat{H}_i = \hat{H}_i(x) = \text{argmax}_{H_j}\{f(x|H_j)P(H_j)\}$$

This is equivalent to the "MAP" rule

$$\hat{H}_i = \text{argmax}_{H_j}\{P(H_j|x)\}$$

REMARKS:

* MAP decision rule minimizes average P_e
* Minimum average P_e is equal to

$$P_e^* = 1 - \sum_{i=1}^{M} E[\phi_i(x)|H_i]P(H_i)$$

* MAP decision rule on x depends only through LR = sufficient statistic
* For equally likely H_i, $P(H_i) = 1/M$ and MAP test is of form

$$\hat{H}_i = \text{argmax}_{H_j}\{f(x|H_j)\}$$

which should be read: "estimate $\hat{H}_i = H_1$ if $f(x|H_1) > f(x|H_0)$." This can be interpreted as the "Maximum likelihood" estimate of true hypothesis H_j

Example 33 *Classifier of Gaussian Means*

* $\underline{X} = [X_1, \ldots, X_n]^T$ are i.i.d. $\mathcal{N}(\mu, \sigma^2)$
* σ^2 is known

OBJECTIVE: classify μ among three possible values

$$\begin{aligned} H_1 : \mu &= \mu_1 \\ H_2 : \mu &= \mu_2 \\ H_3 : \mu &= \mu_3 \end{aligned}$$

Assume equally likely hypotheses

We know that MAP classifier depends on the $\underline{X}$ only through sufficient statistic for μ:

$$\overline{X} = n^{-1} \sum_{i=1}^n X_i$$

which is Gaussian with mean μ and variance σ^2/n.

Therefore, the MAP test is of form:

Decide H_k iff

$$f(\overline{X}|H_k) \geq f(\overline{X}|H_j)$$

where

$$f(\overline{x}|H_k) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left(-\frac{1}{2\sigma^2/n} (\overline{X} - \mu_k)^2\right)$$

or eliminating common factors and taking logarithm

$$\overline{X}\mu_k - \tfrac{1}{2}\mu_k^2 \geq \overline{X}\mu_j - \tfrac{1}{2}\mu_j^2$$

$\Rightarrow$ Linear decision regions!

Concrete example: $\mu_1 = -1$, $\mu_2 = +1$, $\mu_3 = 2$

Plot 3 lines as a function of $\overline{X}$ to find the decision regions:

$$\begin{aligned} \mathcal{X}_1 &= \{\underline{X} : \overline{X} \leq 0\} \\ \mathcal{X}_2 &= \{\underline{X} : 0 < \overline{X} \leq 3/2\} \\ \mathcal{X}_3 &= \{\underline{X} : \overline{X} \geq 3/2\} \end{aligned}$$

These are regions separated by hyperplanes in $\mathcal{X} = \mathbb{R}^n$.

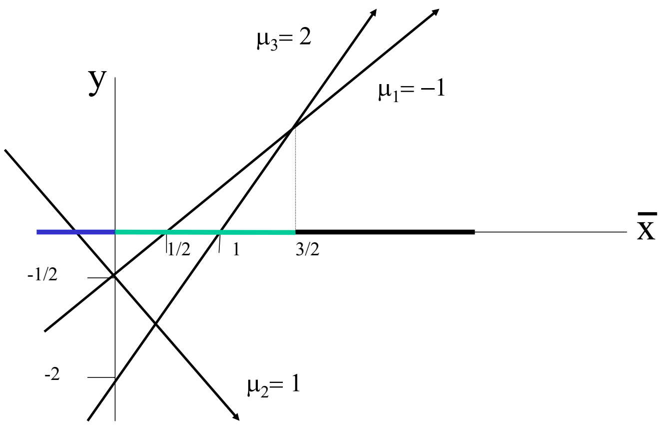


Figure 78: *For three hypotheses on Gaussian mean the decision regions are specified by intersections of three lines* $y = \overline{x}\mu_k - \frac{1}{2}\mu_k^2$, $k = 1, 2, 3$, *which are graphed here over domain* $\overline{x}$.

7.3.3 DEFICIENCIES OF BAYES APPROACH

* Requires assigning prior to θ, H_0, $H_1, \ldots$

* Only ensures best average performance w.r.t. selected prior

* Provides no guaranteed protection against FA, M

7.4 FREQUENTIST APPROACH TO DETECTION

The frequentist approach assumes no priors on H_0 or H_1 so one cannot sensibly define an average probability of error or risk to minimize. Thus we adopt the alternative criterion: constrain FA and minimize M probabilities. It turns out that to find an optimum test satisfying such a constraint we will need to extend our previous definition of a test function ϕ so as to allow for randomized decisions

$$\phi(x) = \begin{cases} 1, & \text{say } H_1 \\ q, & \text{flip a coin w/ prob Heads } (H_1) = q \\ 0, & \text{say } H_0 \end{cases}$$

Note, we have interpretation:

$$\phi(x) = P(\text{say } H_1 | \text{observe } x)$$

False alarm probability and detection probability are functions of θ

$$E_\theta[\phi] = \int_{\mathcal{X}} \phi(x) f(x;\theta) dx = \begin{cases} P_F(\theta), & \theta \in \Theta_0 \\ P_D(\theta), & \theta \in \Theta_1 \end{cases}$$

Definition: A test ϕ is said to be of (FA) level $\alpha \in [0, 1]$ if

$$\max_{\theta \in \Theta_0} P_F(\theta) \leq \alpha$$

Definition: The **power function**' of a test ϕ is

$$\beta(\theta) = P_D(\theta) = 1 - P_M(\theta), \quad \theta \in \Theta_1$$

7.4.1 CASE OF SIMPLE HYPOTHESES: $\theta \in \{\theta_0, \theta_1\}$

$$H_0 : X \sim f(x; \theta_0)$$

$$H_1 : X \sim f(x; \theta_1)$$

Neyman-Pearson Strategy: find most powerful (MP) test ϕ^* of level α:

$$E_{\theta_1}[\phi^*] \geq E_{\theta_1}[\phi]$$

for any other test satisfying $E_{\theta_0}[\phi] \leq \alpha$.

Lemma 1 Neyman Pearson Lemma: *The MP test of level $\alpha \in [0,1]$ is a randomized LRT of the form*

$$\phi^*(x) = \begin{cases} 1, & f(x;\theta_1) > \eta f(x;\theta_0) \\ q, & f(x;\theta_1) = \eta f(x;\theta_0) \\ 0, & f(x;\theta_1) < \eta f(x;\theta_0) \end{cases} \tag{100}$$

where η and q are selected to satisfy

$$E_{\theta_0}[\phi^*] = \alpha$$

Proof 1 of NPL: uses Kuhn-Tucker theory [42] of constrained maximization. If you do not have the background don't worry, we give a more elementary (but longer) proof below.

The MP test maximizes power $E_{\theta_1}[\phi(x)]$ subject to constraint $E_{\theta_0}[\phi(x)] \leq \alpha$. This constrained estimation problem is equivalent to maximizing the unconstrained objective function

$$L(\phi) = E_{\theta_1}[\phi(x)] + \lambda\left(\alpha - E_{\theta_0}[\phi(x)]\right)$$

where $\lambda > 0$ is Lagrange multiplier selected so that solution ϕ^* meets the equality in the original constraint, i.e., $E_{\theta_0}[\phi] = \alpha$.

Now the power can be expressed via the likelihood ratio transformation for expectation, also known as the "Girsanov representation:"

$$E_{\theta_1}[\phi(x)] = E_{\theta_0}\left[\phi(x)\frac{f(x;\theta_1)}{f(x;\theta_0)}\right]$$

and hence:

$$L(\phi) = E_{\theta_0}\left[\phi(x)\left(\frac{f(x;\theta_1)}{f(x;\theta_0)} - \lambda\right)\right] + \lambda\alpha.$$

Our now familiar exchange argument establishes that for a given $x \in \mathcal{X}$ we should choose to assign $\phi(x) = 1$ only if the likelihood ratio exceeds λ. If the LR is less than λ assign $\phi(x) = 0$. This leaves the case for which the LR is equal to λ at which point we randomize the decision, i.e. choose $\phi(x) = q$, $0 < q < 1$, in order to achieve the desired false alarm level. Thus we obtain the randomized LRT (100)of the NPL. $\diamond$

Proof 2 of NPL: more elementary

Need show that for ϕ arbitrary, ϕ^* satisfies

$$E_{\theta_1}[\phi^*] \geq E_{\theta_1}[\phi], \quad \text{when} \quad E_{\theta_0}[\phi^*] = \alpha, \;\; E_{\theta_0}[\phi] \leq \alpha$$

Two steps:

Step 1: Show by enumerating all possible cases of $>, <$ and $=$ between the terms on RHS and LHS

$$\phi^*(x)[f(x;\theta_1) - \eta f(x;\theta_0)] \geq \phi(x)[f(x;\theta_1) - \eta f(x;\theta_0)] \tag{101}$$

Step 2: integrate (101) over all x

$$\int_{\mathcal{X}} \phi^*(x)[f(x;\theta_1) - \eta f(x;\theta_0)]dx \geq \int_{\mathcal{X}} \phi(x)[f(x;\theta_1) - \eta f(x;\theta_0)]dx$$

$$= \underbrace{\int_{\mathcal{X}} \phi^*(x)f(x;\theta_1)dx}_{E_{\theta_1}[\phi^*]} - \eta \underbrace{\int_{\mathcal{X}} \phi^*(x)f(x;\theta_0)dx}_{E_{\theta_0}[\phi^*]}$$

$$\geq \underbrace{\int_{\mathcal{X}} \phi(x)f(x;\theta_1)dx}_{E_{\theta_1}[\phi]} - \eta \underbrace{\int_{\mathcal{X}} \phi(x)f(x;\theta_0)dx}_{E_{\theta_0}[\phi]}$$

Hence

$$E_{\theta_1}[\phi^*] - E_{\theta_1}[\phi] \geq \eta(\underbrace{E_{\theta_0}[\phi^*]}_{=\alpha} - \underbrace{E_{\theta_0}[\phi]}_{\leq\alpha}) \geq 0$$

Which establishes NPL. $\diamond$

RESOURCE ALLOCATION INTERPRETATION OF MP TEST

Assume that you knew the current and future values of a certain set of securities (stocks) x in which you had an opportunity to invest in. You can only buy a single share of each stock. Identify:

$f(x;\theta_0)$ = current value of security x

$f(x;\theta_1)$ = future value of security x

$\phi(x)$ = decision whether or not to invest in security x

α = total available dollars for investment

β = total future value of investment

The NPL says simply: it is best to invest your α\$ in the securities which have the overall highest returns $f(x;\theta_1)/f(x;\theta_0)$. In particular, to maximize the average return you should order all of the stocks in decreasing order of return and start buying stocks in that order until you almost run out of money. At that point flip a biased coin and if it comes up heads, borrow some money from a friend, and buy the next stock on the list. If you choose the right bias on your coin flip you will maximize your expected return and (on average) can pay off the loan to your friend without going into debt (assuming that your friend does not charge interest)!

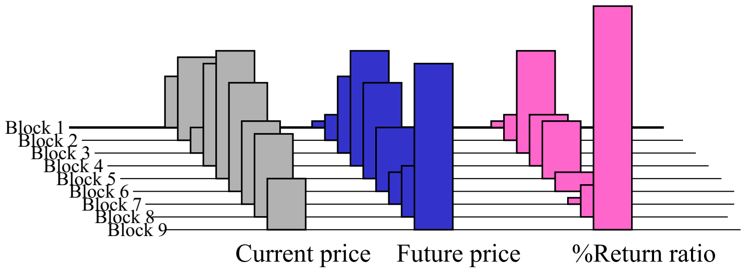


Figure 79: *Future value, current value, and relative return of a set of securities $\mathcal{X}$*

GENERAL REMARKS CONCERNING MP TESTS

Remark 1. shorthand LRT notation

$$\Lambda(x) = f(x;\theta_1)/f(x;\theta_0) \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

Remark 2. P_F of MP test is (Λ denotes $\Lambda(X)$)

$$P_F = E_{\theta_0}[\phi^*(x)] = \underbrace{P_{\theta_0}(\Lambda > \eta)}_{1-F_\Lambda(\eta|H_0)} + qP_{\theta_0}(\Lambda = \eta). \qquad (102)$$

Randomization must be performed only if it is impossible to find an η such $P_{\theta_0}(\Lambda > \eta) = \alpha$. This can only occur if the CDF $F_\Lambda(t|H_0)$ has jump discontinuities, i.e., there exist points $t > 0$ where

$P_{\theta_o}(\Lambda = t) > 0$ and $\Lambda = \Lambda(x)$ is not a cts random variable Otherwise q can be set to zero and randomization is not necessary.

When one cannot find a suitable η that gives $P_{\theta_0}(\Lambda > \eta) = \alpha$, the design procedure is as follows (See Fig. 80):

1. Find the smallest value of t for which $P_{\theta_0}(\Lambda > t)$ is less than α - when there is a jump discontinuity in the CDF this always exists since all CDFs are right continuous. Call this value, α^- and set the threshold η to this value of t.

2. Define $\alpha^+ = P_{\theta_0}(\Lambda = \eta) + \alpha^-$, where α^- and η are determined in step 1. Then from (102) for any value q the test will have the false alarm rate

$$P_F = \alpha^- + q(\alpha^+ - \alpha^-).$$

Setting $P_F = \alpha$ this equation can be solved for q yielding

$$q = \frac{\alpha - \alpha^-}{\alpha^+ - \alpha^-}. \tag{103}$$

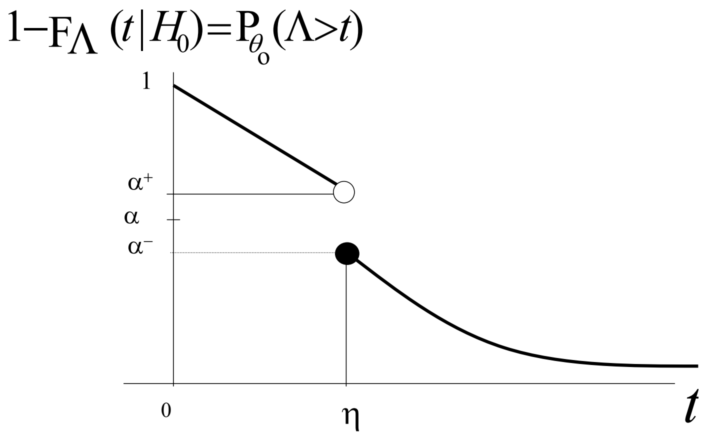


Figure 80: *Randomization is necessary to attain a level α when $1-\alpha$ is not in the range of values of the cdf of Λ.*

Remark 3. LR is identical to Bayes LR for simple hypotheses

Remark 4. Unlike BLRT threshold η is specified by only one quantity α.

Remark 5. If $T = T(X)$ is a sufficient statistic for θ, LRT depends on X only through $T(X)$

Indeed if $f(X;\theta) = g(T,\theta)h(X)$ then

$$\Lambda(X) = g(T,\theta_1)/g(T,\theta_0) = \Lambda(T)$$

Conclude: can formulate the LRT based on p.d.f. of T instead of the p.d.f. of entire data sample X.

7.5 ROC CURVES FOR THRESHOLD TESTS

All threshold tests have P_F and P_D indexed by a parameter η.

The Receiver Operating Characteristic (ROC) is simply the plot of the parametric curve $\{P_F(\eta, q), P_D(\eta, q))\}_{\eta,q}$.

Equivalently, ROC is the plot of $\beta = P_D$ vs $\alpha = P_F$.

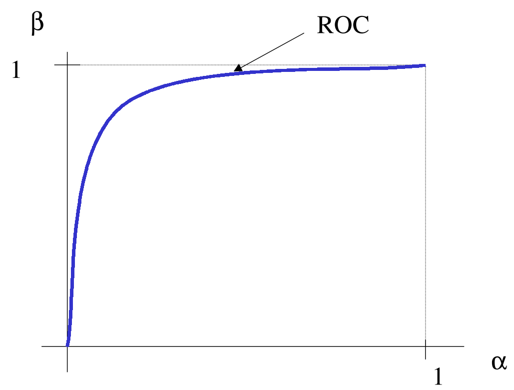


Figure 81: *A typical ROC curve.*

PROPERTIES OF ROC'S

1. ROC for coin flip detector ($\phi(x) = q$ independent of data) is a diagonal line with slope =1

$$\begin{aligned} \alpha = P_F &= E_{\theta_0}[\phi] = q \\ \beta = P_D &= E_{\theta_1}[\phi] = q \end{aligned}$$

2. ROC of any MP test always lies above diagonal: MP test is "unbiased" test

Definition: a test ϕ is unbiased if its detection probability β is at least as great as its false alarm α: $\beta \geq \alpha$.

3. ROC of any MP test is always convex cap (concave).

To see concavity, let (α_1, β_1) be the level and power of a test ϕ_1 and (α_2, β_2) be the level and power of a test ϕ_2. Define the test

$$\phi_{12} = p\phi_1 + (1-p)\phi_2$$

This test can be implemented by selecting ϕ_1 and ϕ_2 at random with probability p and $1-p$, respectively. The level of this test is

$$\alpha_{12} = E_0[\phi_{12}] = pE_0[\phi_1] + (1-p)E_0[\phi_2] = p\alpha_1 + (1-p)\alpha_2$$

and its power is similarly

$$\beta_{12} = E_1[\phi_{12}] = p\beta_1 + (1-p)\beta_2$$

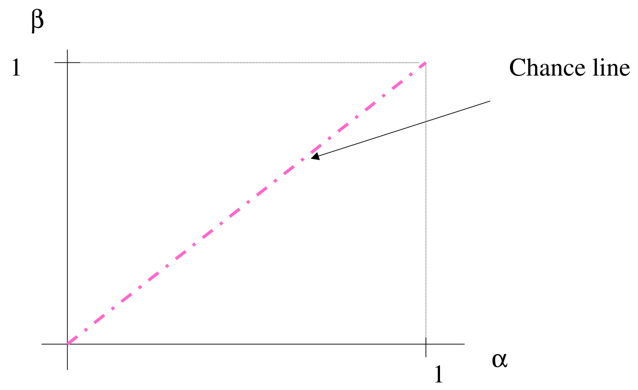


Figure 82: *ROC curve for coin flip detector.*

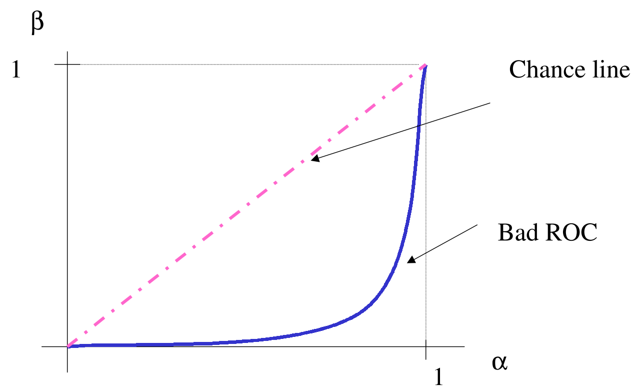


Figure 83: *ROC curve for MP test always lies above diagonal.*

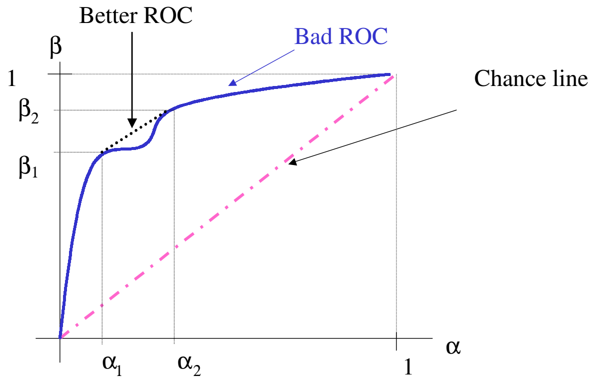


Figure 84: *ROC of any MP test is always convex cap. A test with non-convex ROC (thick line) can always be improved by randomization which has effect of connecting two endpoints ((α_1, β_1) and (α_2, β_2) on ROC by straight line.*

Thus, as p varies between 0 and 1, ϕ_{12} has performance $(\alpha_{12}, \beta_{12})$ which varies on a straight line connecting the points (α_1, β_1) and (α_2, β_2).

4. If ROC curve is differentiable, MP-LRT threshold needed for attaining any pair $(\alpha, P_D(\alpha))$ on ROC can be found graphically as slope of ROC at the point α.

$$\eta = \frac{d}{d\alpha} P_D(\alpha)$$

5. When the hypotheses H_0 and H_1 are simple, the MP-LRT threshold that attains minmax P_e can also be found graphically by intersection of line $P_M = 1 - P_D = P_F$ and ROC.

Example 34 *Test against uniform density*

Two hypotheses on a scalar r.v. x

$$H_0 : f(x) = f_0(x)$$

$$H_1 : f(x) = f_1(x)$$

where f_0 and f_1 are two densities shown in Fig. 87.

Objective: find the MP-LRT

Solution: LRT is

$$\Lambda(x) = \frac{f_1(x)}{f_0(x)} \overset{H_1}{\underset{H_0}{\gtrless}} \eta$$

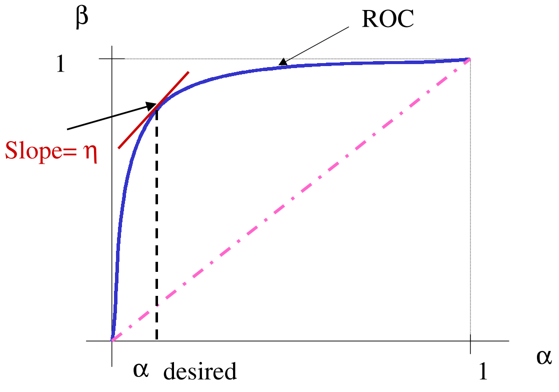


Figure 85: *Threshold of MP-LRT can be found by differentiation of ROC curve.*

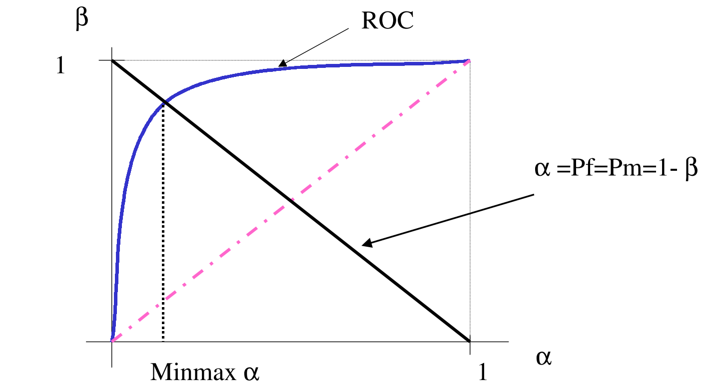


Figure 86: *Threshold of min-max Bayes test can be found by intersection.*

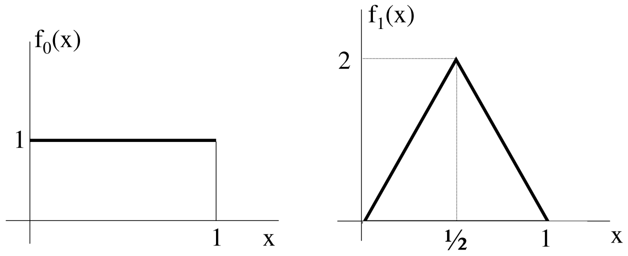


Figure 87: *Two densities to be tested*

or equivalently

$$f_1(x) \overset{H_1}{\underset{H_0}{\gtrless}} \eta f_0(x)$$

From Fi. 88 it is obvious that for a given η the H_1 decision region is

$$\mathcal{X}_1 = \begin{cases} \{\eta/4 < x < 1 - \eta/4\}, & 0 \leq \eta \leq 2 \\ \text{empty}, & o.w. \end{cases}$$

Setting threshold

Select η to meet constraint $P_F = \alpha$.

FIRST: attempt to set η without randomization ($q = 0$).

Assume $\eta \in [0, 2]$

$$\begin{aligned} \alpha &= P(X \in \mathcal{X}_1 | H_0) = \int_{\eta/4}^{1-\eta/4} f_0(x) dx \\ &= 1 - \eta/2 \end{aligned}$$

Hence required η is simply

$$\eta = 2(1 - \alpha)$$

and we see that no randomization is required.

Power of MP-LRT is:

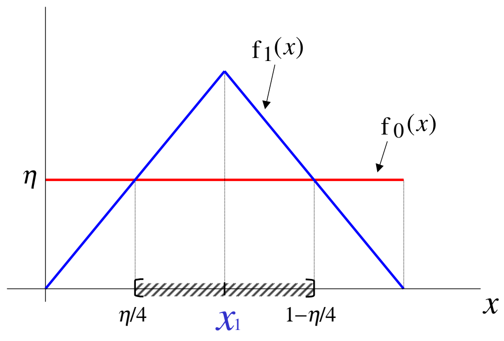


Figure 88: *Region $\mathcal{X}_1$ for which MP-LRT decides H_1 are set of values x for which triangle exceeds horizontal line of height η.*

$$\begin{aligned} P_D &= P(X \in \mathcal{X}_1 | H_1) = \int_{\eta/4}^{1-\eta/4} f_1(x)dx \\ &= 2\int_{\eta/4}^{\frac{1}{2}} f_1(x)dx = 2\int_{\eta/4}^{\frac{1}{2}} 4x dx \\ &= 1 - \eta^2/4 \end{aligned}$$

Plug in level α threshold $\eta = 2(1-\alpha)$ to power expression to obtain the ROC curve

$$\beta = 1 - (1-\alpha)^2$$

Example 35 *Detecting an increase in Poisson rate*

Let X be the reading of the number of photons collected by a charge coupled device (CCD) array over a certain period of time. In ambient conditions the average number of photons incident on the array is fixed and known, let's call it θ_0. This is sometimes called the dark current rate [38]. When a known source of photons is present the photon rate increases to a known value θ_1 where $\theta_1 > \theta_0$. The goal of the photodetector is to detect the presence of the source based on measuring $X = x$. It is customary to assume that X is a Poisson random variable

$$X \sim f(x;\theta) = \frac{\theta^x}{x!}e^{-\theta}, \quad x = 0, 1, \ldots$$

and the problem is to detect the increase from θ_0 to θ_1 in the Poisson rate parameter θ, i.e., to test the simple hypotheses

$$\begin{aligned} H_0 : \theta &= \theta_0 \\ H_1 : \theta &= \theta_1 \end{aligned}$$

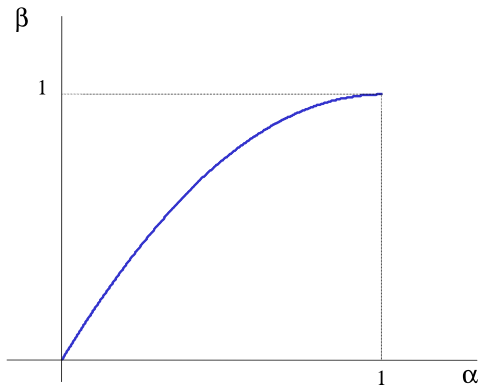


Figure 89: *ROC curve for uniform vs. triangle pdf example.*

where $\theta_1 > \theta_0 > 0$. Here we consider the design of a MP test of prescribed level $\alpha \in [0, 1]$.

Solution: we know that the MP test is a LRT

$$\Lambda(x) = \left(\frac{\theta_1}{\theta_0}\right)^x e^{\theta_0 - \theta_1} \overset{H_1}{\underset{H_0}{\gtrless}} \eta.$$

Since the logarithm is a monotone increasing function, and $\theta_1 > \theta_0$, the MP-LRT is equivalent to a linear test

$$x \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

where (needed for Bayes LRT but not for MP-LRT) $\gamma = \frac{\ln \eta + \theta_1 - \theta_0}{\ln(\theta_1/\theta_0)}$.

We first try to set threshold γ without randomization:

$$\alpha = P_{\theta_0}(X > \gamma) = 1 - \mathcal{P}\mathrm{o}_{\theta_0}(\gamma)$$

where $\mathcal{P}\mathrm{o}_\theta(\cdot)$ is the CDF of a Poisson r.v. with rate θ. Here we run into a difficulty illustrated by Fig. 90. As the Poisson CDF is not continuous only a discrete number of values are attainable by the nonrandomized LRT

$$\alpha \in \{\alpha_i\}_{i=1}^\infty, \qquad \alpha_i = 1 - \mathcal{P}\mathrm{o}_{\theta_0}(i).$$

Assume $\alpha \in (\alpha_i, \alpha_{i+1})$. Then we need to randomize the LRT by selecting γ, q to satisfy:

$$\alpha = P_{\theta_0}(X > \gamma) + q P_{\theta_0}(X = \gamma).$$

Following the procedure described in connection with equation (103) we select

$$\gamma = \gamma^* := \mathcal{P}\mathrm{o}_{\theta_0}^{-1}(1 - \alpha_i)$$

which gives $P_{\theta_0}(X > \gamma^*) = \alpha_i$, and we set the randomization according to

$$q = q^* := \frac{\alpha - \alpha_i}{\alpha_{i+1} - \alpha_i}.$$

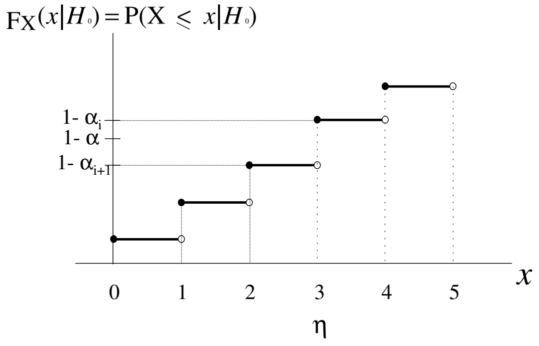


Figure 90: *CDF of LR test statistic for testing increase in Poisson rate is staircase function*

With these settings the power of the randomized MP-LRT is simply

$$P_D = P_{\theta_1}(X > \gamma^*) + q^* P_{\theta_1}(X = \gamma^*),$$

which is plotted as an ROC curve in Fig. 91.

Example 36 *On Off keying (OOK) in Gaussian noise*

On-off keying is a type of binary modulation that is used in many digital communications systems and can be traced back to the early days of Morse code and the telegraph. Over a single bit interval the integrated output X of the receiver can be modeled as either noise alone W (if the transmitted bit is zero) or a constant, assumed equal to 1, plus noise. The decoder has to decide between

$$\begin{aligned} H_0 : X &= W \\ H_1 : X &= 1 + W \end{aligned}$$

where we assume $W \sim \mathcal{N}_1(0,1)$, i.e., the received SNR is 0dB. The LR statistic is simply expressed as

$$\Lambda(x) = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-1)^2}}{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}} = e^{x-\frac{1}{2}}.$$

As usual the ROC curve is obtained from $P_D = P(X > \lambda|H_1) = P(X - 1 > \lambda - 1|H_1) = 1 - \mathcal{N}(\lambda - 1)$. Substituting the above λ expressions into this equation

$$\beta = P_D = 1 - \mathcal{N}(\lambda - 1) = 1 - \mathcal{N}(\mathcal{N}^{-1}(1 - \alpha) - 1).$$

This curve is shown in Fig. 93 along with operating points for three different ways of setting the threshold of the LRT: the Bayes LRT, the minmax LRT, and the MP-LRT.

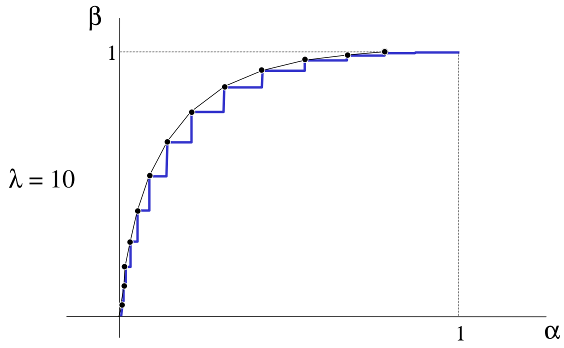


Figure 91: *Power curves of LRT for detecting an increase in rate of a Poisson r.v. The smooth curve is the (randomized) MP test while the staircase curve is the non-randomized LRT.*

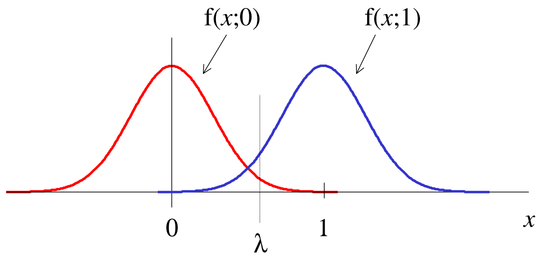


Figure 92: *Densities under H_0 and H_1 for on-off keying detection.*

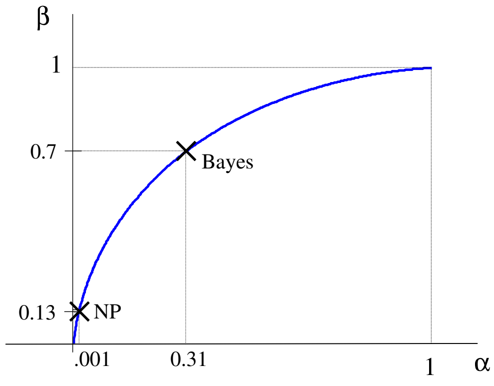


Figure 93: *ROC curve for Gaussian OOK example.*

1. Min P_e (Bayes) test for equally likely H_0, H_1 ($\eta = 1$):

$$x \underset{H_0}{\overset{H_1}{\gtrless}} \ln \eta + \tfrac{1}{2} = \tfrac{1}{2}$$

2. Minmax test:

$$x \underset{H_0}{\overset{H_1}{\gtrless}} \ln \eta + \tfrac{1}{2} := \lambda,$$

where λ is chosen to satisfy

$$P_F = 1 - \mathcal{N}(\lambda) = \mathcal{N}(\lambda - 1) = P_M.$$

The solution to this equation is again $\lambda = \frac{1}{2}$ since $\mathcal{N}(-x) = 1 - \mathcal{N}(x)$.

3. MP test of level α:

$$x \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

where $\alpha = P(X > \lambda | H_0) = 1 - \mathcal{N}(\lambda)$ or

$$\lambda = \mathcal{N}^{-1}(1 - \alpha)$$

A quantitative performance comparison is shown in Table 7.5 where we have specified the FA level $\alpha = 0.001$ for the MP-LRT (corresponding MP-LRT threshold is $\lambda = 2.329$).

Note from the table that the Bayes and minimax test have identical performance since they use identical threshold and that the NP test has much lower FA rate but also significantly lower P_D and higher P_e.

	P_F	P_D	P_e
Bayes	0.31	0.69	0.31
Minmax	0.31	0.69	0.31
NP	0.001	0.092	0.5

Table 1: Performance comparisons for three different threshold settings in OOK example.

7.6 BACKGROUND AND REFERENCES

There are many good textbooks on detection theory for signal processing, control and communications. The books by Van Trees [73] and Whalen [76] are classics in the field. One of the earliest relevant reference books covering signal detection theory is Middleton's statistical communications theory opus [46] which adopts a mathematical-physics perspective. The more recent book by Helstrom [25] takes a similar but more engineering-oriented perspective. Another recent book with a signal processing focus on signal detection is Srinath, Rajasekaran and Viswanathan [67]. For a somewhat more advanced mathematical treatment the reader may wish to consult the book by Poor [55]. The above books concentrate on continuous time measurements which we do not cover in this chapter. The book by Mood, Graybill and Boes [48] has a very nice but elementary treatment of the statistical methodology. More advanced treatments are found in books by Bickel and Doksum [7], Lehmann [39] and Ferguson [16].

7.7 EXERCISES

7.1 A proprietary binary hypothesis test ϕ is implemented in a software package which you are considering purchasing based on a trial examination period. You run several experiments and obtain the following table of probabilities of detection β vs. false alarm α

α	β
0.1	0.2
0.3	0.4
0.5	0.8
0.7	0.9

Comment on the quality of this test. Could you improve on this test? If so specify the improved test and compute its ROC curve.

7.2 Let Z be a random variable with values in the interval $[-1, 1]$ having density function

$$p_\theta(z) = \frac{1}{2}\frac{3}{3+\theta}\left(\theta z^2 + 1\right)$$

where $\theta > 0$. Note θ controls the deviation of p_θ from the uniform density p_0. You are to test Z against non-uniformity given a *single sample* Z.

(a) Assuming priors $p = P(H_1) = 1 - P(H_0)$ (note this is opposite to the convention of this chapter) derive the minimum probability of error (MAP) test for the simple hypotheses $H_0 : \theta = 0$ vs. $H_1 : \theta = \theta_1$, where θ_1 is a fixed and known positive value.

(b) Find an expression for the ROC curve and plot for $\theta_1 = 0, 1, 10$.

(c) Now find the form of the min-max test. Show how you can use your answer to part b) to graphically determine the min-max threshold.

(d) Derive the MP test for the same simple hypotheses as in part (a).

7.3 Let Z be a single observation having density function

$$p_\theta(z) = (2\theta z + 1 - \theta), \qquad 0 \leq z \leq 1$$

where $-1 \leq \theta \leq 1$.

(a) Find the most powerful test between the hypotheses

$$\begin{aligned} H_0 &: \theta = 0 \\ H_1 &: \theta = 1 \end{aligned}$$

Be sure to express your test in terms of the false alarm level $\alpha \in [0, 1]$. Plot the ROC curve for this test.

(b) repeat part (a) for $H_1 : \theta = -1$.

7.4 It is desired to test the following hypotheses based on a single sample x:

$$\begin{aligned} H_0 &: x \sim f_0(x) = \frac{3}{2}\, x^2, -1 \leq x \leq 1 \\ H_1 &: x \sim f_1(x) = \frac{3}{4}\, (1 - x^2), -1 \leq x \leq 1 \end{aligned}$$

(a) Under the assumption that the prior probabilities of H_0 and H_1 are identical, find the minimum probability of error (Bayes) test.

(b) Find the Most Powerful test of level $\alpha \in [0, 1]$.

(c) Derive and plot the ROC curve for these tests.

7.5 Let $f(x|H_0)$ and $f(x|H_1)$ be densities of an observed r.v. x and assume that the likelihood ratio $\Lambda = f(x|H_1)/f(x|H_0)$ has corresponding densities $f_\Lambda(\lambda|H_0)$ and $f_\Lambda(\lambda|H_1)$ under H_0 and H_1, respectively. Show that the slope $d\beta/d\alpha$ at a point α of the ROC of the LRT is equal to the threshold η attaining level α. (Hint: show that $d\beta/d\alpha = f_\Lambda(\eta|H_1)/f_\Lambda(\eta|H_0)$ and then apply $f_\Lambda(u|H_k) = \int_{\{x:\Lambda(x)=u\}} f(x|H_k)dx$, $k = 0, 1$.)

7.6 Let a detector have the ROC curve $\{(\alpha, \beta) : \alpha \in [0, 1]\}$ where the power function $\beta = \beta(\alpha)$ is a function of the false alarm level α. The area under the ROC is defined as

$$\text{AUC} = \int_0^1 \beta(\alpha)d\alpha$$

The AUC is frequently used as an alternative to the power function to assess the performance of various detectors. Assume simple hypotheses and invoke properties of ROCs in answering the following questions.

(a) Show that among all tests the MP LRT maximizes AUC.

(b) Show the following inequalities for the AUC of a MP LRT

$$\frac{1}{2} \leq \text{AUC} \leq \beta(\tfrac{1}{2}) \leq 1$$

(c) Show that for any LRT whose ROC $\beta(\alpha)$ is differentiable in α

$$\text{AUC} = 1 - \int_0^1 \alpha\eta(\alpha)d\alpha$$

where $\eta = \eta(\alpha)$ is the LRT's threshold attaining the false alarm level α. When combined with (b) this implies the interesting result for LRT's: as the integral is bounded $\lim_{\alpha\to 0}(\alpha\eta) = 0$, i.e. α decreases to zero faster than $\eta(\alpha)$ increases to ∞.

7.7 Available is a single random sample X from density $f_\theta(x)$, where $\theta \in \{0,1\}$ and

$$\begin{aligned} f_1(x) &= \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) \\ f_0(x) &= \frac{1}{2} \exp(-|x|). \end{aligned}$$

You are to develop tests for the hypotheses

$$\begin{aligned} H_0 &: \theta = 0 \\ H_1 &: \theta = 1 \end{aligned}$$

(a) Derive the (non-randomized) likelihood ratio test (LRT) and the induced decision region $\mathcal{X}_1$={x : decide H_1} for given threshold η. Draw the decision region as a function of the threshold, i.e. plot the region $\mathcal{X}_1$ for several values of $\eta > 0$.

(b) Compute the false alarm probability P_F and the detection probability P_D of the test. Find an equation for and plot the ROC curve.

(c) Find the optimal Bayes test when H_0 and H_1 have prior probabilities $P(H_0) = 1/4$ and $P(H_1) = 3/4$, the cost of correct decisions is zero and the cost of incorrect decisions is one. What is P_F and P_D for this test?

(d) Find the optimal minimax test for unknown $P(H_0), P(H_1)$. What is P_F and P_D for this test?

(e) Find the Most Powerful test of level $\alpha = 0.2$. What is P_F and P_D for this test?

7.7 Here we consider the problem of simultaneously testing hypotheses on a large number of independent variables, the so called problem of multiple comparisons, in the Bayesian setting. Consider a set of N i.i.d. pairs of random variables $\{(X_i, \theta_i)\}_{i=1}^N$. Conditioned on θ_i, X_i has density $f_{\theta_i}(x)$, where $\theta_i \in \{0,1\}$. The θ_i have prior probabilities $P(\theta_i = 1) = p$ and $P(\theta_i = 0) = 1 - p$. Given that we observe the X_i's but not the θ_i's we wish to test the following N hypotheses

$$\begin{aligned} H_0(k) &: \theta_k = 0 \\ H_0(k) &: \theta_k = 1, \end{aligned}$$

or, equivalently, $H_0(k) : X_k \sim f_0$ vs $H_1 : X_k \sim f_1$, $k = 1, \ldots, N$.

(a) Define the test function $\phi_i = \phi(X_i) \in \{0,1\}$ for testing the i-th hypothesis, i.e., if $\phi_i = 1$ we decide that $\theta_i = 1$ else decide $\theta_i = 0$. Show that the false alarm and miss probabilities associated with ϕ_i can be represented as:

$$P_F = E[\phi_i(1-\theta_i)]/(1-p)$$

$$P_M = E[(1-\phi_i)\theta_i]/p,$$

respectively, and that the total probability of error is

$$P_e(i) = E[\phi_i(1-\theta_i)] + E[(1-\phi_i)\theta_i].$$

By using nested conditional expectation on $P_e(i)$ show that the optimal test function that minimizes $P_e(i)$ is the one that assigns $\phi(X_i) = 1$ whenever

$$E[\theta_i|X_i]/(1 - E[\theta_i|X_i]) > 1,$$

and this is equivalent to the MAP decision rule.

(b) For a given set of samples $\{(X_i, \theta_i)\}_{i=1}^N$ the number of declared detections, or "discoveries," is defined as the random variable $M = \sum_{i=1}^N \phi_i$. The Bayesian false discovery rate (FDR) is defined as the average proportion of false positives occurring among these M "discoveries"

$$\mathrm{FDR} = E[\sum_{i=1}^{N} \phi_i (1-\theta_i)/M].$$

Constrain the FDR of these tests to have $\mathrm{FDR} \leq q$. Subject to this constraint we would like to minimize the average number of missed "$\theta_i = 1$" events. Equivalently, we want to maximize the average number of true positives discovered:

$$\mathrm{TP} = E[\sum_{i=1}^{N} \phi_i \theta_i].$$

Similarly to how we derived the Neyman=Pearson MP test in class, this optimal Bayesian FDR constrained test of level q must maximize the Lagrangian

$$\mathcal{L}(\phi_1, \ldots, \phi_N) = \mathrm{TP} + \lambda(q - \mathrm{FDR})$$

where λ is an undetermined multiplier selected to satisfy the FDR constraint. Show that the optimal Bayesian FDR test of level q is the following "MAP test with linearly increasing threshold:" assign $\phi_i = 1$ to all i such that

$$T_i = \frac{P(\theta_i = 0|X_i)}{P(\theta_i = 1|X_i)} < M/\lambda,$$

where $\lambda > 0$ is selected to attain $\mathrm{FDR} = q$. This test can be implemented by rank ordering all of the scores (also called "posterior odds ratios") T_i in increasing order $T_{(1)} \leq \ldots \leq T_{(N)}$ and finding the first index M at which $T_{(i)}$ goes above the straight line $T_i = i/\lambda$. Only those hypotheses having scores less than than M/λ should be declared as valid discoveries.

(c) Let $f_0(x) = a_0 e^{-a_0 x}$ and $f_1(x) = a_1 e^{-a_1 x}$, for $x > 0$ and $a_0 > a_1 > 0$. For $N = 2$ derive an expression for the threshold λ in part (b) and compute the ROC. Use Matlab or other tool to plot the ROC if you like. You might find it interesting to simulate this test for large N.

End of chapter

8 DETECTION STRATEGIES FOR COMPOSITE HYPOTHESES

In practical detection applications it is rare that one fully knows the distributions under each of the hypotheses. When these distributions are approximated as parametric models one can express this uncertainty as an unknown variation of the parameters and formulate the detection problem as a test of composite hypotheses consisting of unknown parameter values under either H_0 or H_1 or both. A Bayes approach to handling such uncertainty would be to assign priors to these parameter values and derive the Bayes-optimal LRT, as discussed in the previous chapter, that minimizes the average probability of decision error. Thus the Bayes strategy is not at all complicated by the presence of composite hypotheses and we need not discuss it further.

However, if one is interested in maximizing detection probability while controlling false alarm the presence of uncertainty in parameters presents challenges. Most powerful tests of a given level are rarely extendible to composite hypotheses; there seldom exists a test which is most powerful at a prescribed level α for all values of the unknown parameters. There are a number of other non-Bayesian strategies that can be adopted and in this chapter we present several of these whose aim is to guarantee performance more or less robust to unknown parameter variations.

We will first present strategies for composite hypothesis testing that have finite sample optimality properties. These include the uniformly most powerful test, the locally most powerful tests, unbiased tests, CFAR tests, and minimax tests. We then present a sub-optimal but very widely adopted strategy called the Generalized Likelihood Ratio (GLR) test which is an LRT implemented with plug-in estimates of the unknown parameters. The GLR test is only briefly introduced in this chapter. Chapters 9 and 12 continue the development of the GLRT in the context of the Gaussian hypothesis testing problem.

A basic property that any reasonable test must have is that its P_D never be less than its P_F for any value of the parameters. If a threshold test violated this property then one could easily beat this test by being a contrarian and deciding H_0 when it decided H_1 and vice-versa. Indeed for simple hypotheses the LRT always has this property (recall our discussion of ROCs in Section 7.5). This property of tests is called *unbaisedness*. Mathematically speaking, a test is said to be unbiased if

$$E_\theta[\phi] \geq \alpha, \qquad \text{all } \theta \in \Theta_1. \tag{104}$$

Otherwise, there will be some θ for which the test gives $P_D < P_F$, and the test is said to be biased.

While unbiasedness is a basic property one expects in a test, a "ideal" test might not only be unbiased but also a most powerful test for any value of $\theta \in \Theta_1$, i.e., a uniformly most powerful (UMP) test. A test that is UMP is always unbiased but when an UMP test does not exist even a LRT test can be biased. This will be illustrated by an example below.

8.1 UNIFORMLY MOST POWERFUL (UMP) TESTS

After reading this section the reader may feel that UMP tests are better described as a miracle than a strategy. However, development of the theory of UMP tests is very instructive as it helps understand the challenges posed in trying to test composite hypotheses. We start by considering a simple null hypothesis with composite alternative

$$H_0 : \theta \;=\; \theta_0 \tag{105}$$
$$H_1 : \theta \;\in\; \Theta_1. \tag{106}$$

Note that the power $P_D = P_D(\theta_1)$ of any good FA level α detector usually varies as a function of $\theta_1 \in \Theta_1$. For example, if θ_1 parameterized signal strength one would expect a good detector to have better P_D as signal strength increased.

Recall from Chapter 7 that a test of (106) is characterized by its test function ϕ (97). This test is of level α if

$$P_F = E_{\theta_0}[\phi] \leq \alpha.$$

A false alarm constrained uniformly most powerful test (UMP) is a test which is MP for any and all values of $\theta \in \Theta_1$, i.e., it is more powerful than any other similarly constrained test (Fig. 94). We give a formal definition below

Definition: a test ϕ^* is a uniformly most powerful (UMP) test of level α if for any other level α test ϕ

$$\beta^*(\theta) = E_\theta[\phi^*] \geq E_\theta[\phi] = \beta(\theta), \quad \text{for all } \theta \in \Theta_1.$$

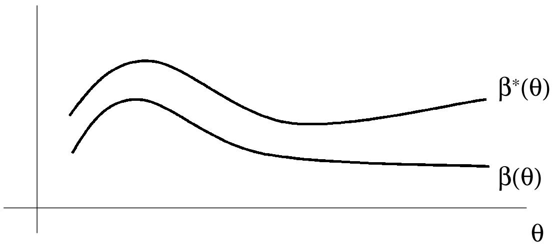


Figure 94: *Power curve $\beta^*(\theta)$, $\theta \in \Theta_1$ of a UMP test is uniformly higher than that of any other test of the same level α.*

There are two steps for discovering a UMP when it exists and, short of this, establishing that a UMP does not exist:

Step 1: Fix $\theta \in \Theta_1$ and find MP test of level α

Step 2: if decision regions of this MP test do not depend on our choice of $\theta \in \Theta_1$ then the MP test is actually UMP over $\theta \in \Theta_1$.

Example 37 *Tests of mean in Gaussian sample with known variance*

$\underline{X} = [X_1, \ldots, X_n]^T$ i.i.d., $X_1 \sim \mathcal{N}(\mu, \sigma^2)$, σ^2 is known.

Three cases of interest:

$$\underbrace{\begin{array}{c} H_0 : \mu = 0 \\ H_1 : \mu > 0 \end{array}}_{\text{Case I}} \quad \underbrace{\begin{array}{c} H_0 : \mu = 0 \\ H_1 : \mu < 0 \end{array}}_{\text{Case II}} \quad \underbrace{\begin{array}{c} H_0 : \mu = 0 \\ H_1 : \mu \neq 0 \end{array}}_{\text{Case III}}$$

Step 1: find LRT for fixed μ under H_1

It suffices to work the problem based on a sufficient statistic $T = \overline{X}$ for μ. We know:

$$\begin{aligned} \overline{X} &\sim \mathcal{N}(0, \sigma^2/n), \quad \text{under } H_0 \\ \overline{X} &\sim \mathcal{N}(\mu, \sigma^2/n), \quad \text{under } H_1 \end{aligned}$$

therefore,

$$\begin{aligned} \Lambda(\mu) &= \frac{f(\overline{X};\mu)}{f(\overline{X};0)} = \frac{\exp\left(-\frac{(\overline{X}-\mu)^2}{2\sigma^2/n}\right)}{\exp\left(-\frac{\overline{X}^2}{2\sigma^2/n}\right)} \\ &= \exp\left(\frac{n\mu}{\sigma^2}\overline{X} - \frac{n\mu^2}{2\sigma^2}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \end{aligned}$$

For clarity, our notation explicitly brings out the dependance of the likelihood ratio on μ. Note that $\Lambda(\mu)$ is monotone increasing in $\mu\overline{X}$ so that one form of the MP-LRT is

$$\mu\left(\frac{\sqrt{n}\,\overline{X}}{\sigma}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

CASE I: Single sided alternative $H_1 : \mu > 0$

In this case μ can be absorbed into RHS without changing inequalities:

$$T(\underline{X}) = \frac{\sqrt{n}\,\overline{X}}{\sigma} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma^+$$

or equivalently, MP-LRT is the linear detector

$$\sum_{i=1}^n X_i \underset{H_0}{\overset{H_1}{\gtrless}} \gamma' = \gamma^+ \sqrt{n}\,\sigma^2$$

Next we must set threshold:

Since we know $\overline{X} \sim \mathcal{N}(0, \sigma^2/n)$ under H_0:

$$\alpha = P_0(\underbrace{\sqrt{n}\,\overline{X}/\sigma}_{\mathcal{N}(0,1)} > \gamma^+) \;=\; 1 - \mathcal{N}(\gamma^+)$$

Or

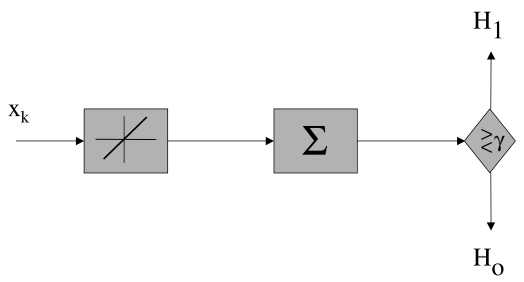


Figure 95: *Optimal detector for positive Gaussian mean is a memoryless linear device followed by a summer and decision mechanism.*

$$\gamma^+ = \mathcal{N}^{-1}(1-\alpha)$$

Final form of MP-LRT for $H_1 : \mu > 0$ reveals that it is UMP against unknown positive μ

$$y \stackrel{\text{def}}{=} \frac{\sqrt{n}\,\overline{X}}{\sigma} \underset{H_0}{\overset{H_1}{\gtrless}} \mathcal{N}^{-1}(1-\alpha)$$

Equivalent form in terms of sample mean statistic

$$\overline{X} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sigma}{\sqrt{n}}\mathcal{N}^{-1}(1-\alpha)$$

Equivalent form in terms of sum statistic

$$\sum_{i=1}^n X_i \underset{H_0}{\overset{H_1}{\gtrless}} \sqrt{n}\,\sigma\,\mathcal{N}^{-1}(1-\alpha)$$

Power of single sided test:
Since $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$ under H_1

$$\begin{aligned}\beta &= P_1(\underbrace{\sqrt{n}\,\overline{X}/\sigma}_{\mathcal{N}(\sqrt{n}\,\mu/\sigma,\,1)} > \gamma^+) \\ &= 1 - \mathcal{N}\left(\gamma^+ - \frac{\sqrt{n}\mu}{\sigma}\right)\end{aligned}$$

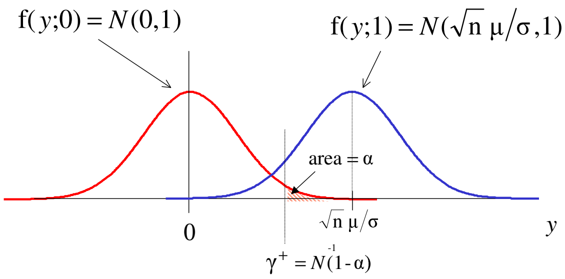


Figure 96: *Threshold γ^+ of MP-LRT for $H_0 : \mu = 0$ vs. $H_1 : \mu > 0$ in i.i.d. Gaussian with known variance. $f(y;0)$ and $f(y;1)$ denote the densities of $y = \frac{\sqrt{n}\,\overline{X}}{\sigma}$ under H_0 and H_1, respectively.*

$$= \quad 1 - \mathcal{N}\left(\mathcal{N}^{-1}(1-\alpha) - d\right)$$

where d is the **positive** *detectability index*

$$d \quad = \quad \frac{\sqrt{n}\mu}{\sigma}$$

$$= \quad \frac{|E[T|H_1] - E[T|H_0]|}{\sqrt{\mathrm{var}_0(T)}}$$

CASE II: Single sided alternative $H_1 : \mu < 0$

Recall that the MP LRT for fixed μ has the form

$$\mu \, \frac{\sqrt{n}\,\overline{X}}{\sigma} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \gamma$$

This is now equivalent to

$$\frac{\sqrt{n}\,\overline{X}}{\sigma} \begin{array}{c} H_0 \\ > \\ < \\ H_1 \end{array} \gamma^-$$

Setting threshold:

$$\alpha = P_0(\underbrace{\sqrt{n}\,\overline{X}/\sigma}_{\mathcal{N}(0,1)} \leq \gamma^-) \quad = \quad \mathcal{N}(\gamma^-)$$

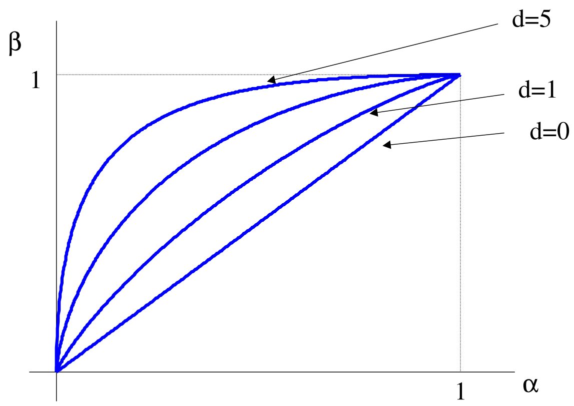


Figure 97: *The ROC curve of MP-LRT for $H_0 : \mu = 0$ vs. $H_1 : \mu > 0$ for n i.i.d. Gaussian with known variance for various values of d.*

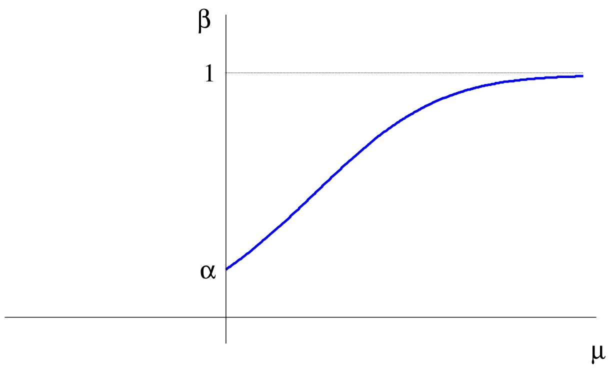


Figure 98: *The power curve of MP-LRT for $H_0 : \mu = 0$ vs. $H_1 : \mu > 0$ for n i.i.d. Gaussian with known variance plotted as a function of $d > 0$*

or now

$$\gamma^- = \mathcal{N}^{-1}(\alpha) = -\mathcal{N}^{-1}(1-\alpha)$$

Again we see MP-LRT is UMP against unknown negative μ

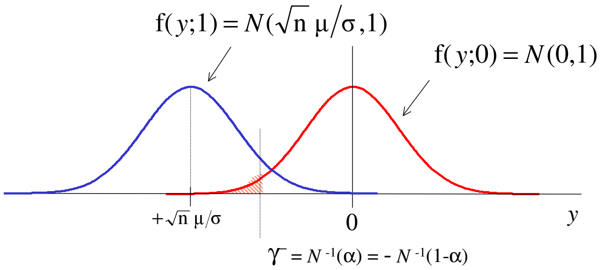


Figure 99: *Threshold determination for MP-LRT of $H_0 : \mu = 0$ vs. $H_1 : \mu < 0$ for n i.i.d. Gaussian observations with known variance*

Power curve for $\mu < 0$ case can be derived similarly

$$\beta = 1 - \mathcal{N}\left(\mathcal{N}^{-1}(1-\alpha) + \underbrace{\frac{\sqrt{n}\,\mu}{\sigma}}_{-|d|}\right)$$

where d is now negative valued

$$d = \frac{\sqrt{n}\mu}{\sigma}$$

CASE III: Double sided alternative $H_1 : \mu \neq 0$

Recall again the form of MP LRT for fixed μ

$$\mu\,\frac{\sqrt{n}\,\overline{X}}{\sigma} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

Unfortunately it is no longer possible to absorb μ into threshold without affecting the inequalities. We thus conclude that the decision region varies depending on sign of μ. Therefore no UMP test exists.

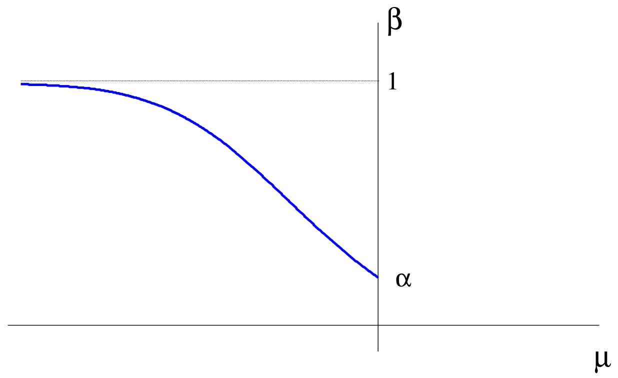


Figure 100: *The power curve of MP-LRT for* $H_0 : \mu = 0$ *vs.* $H_1 : \mu < 0$ *in i.i.d. Gaussian with known variance plotted as a function of* d

If we use single sided test from CASE I then

$$\beta = 1 - \mathcal{N}\left(\mathcal{N}^{-1}(1-\alpha) - d\right),$$

which means that $P_D < P_F$ and the test is biased for $d < 0$, i.e., $\mu < 0$. On the other hand if we use single sided test from CASE II then

$$\beta = 1 - \mathcal{N}\left(\mathcal{N}^{-1}(1-\alpha) + d\right),$$

which is biased for $d > 0$, i.e., $\mu > 0$.

Example 38 *Test of variance in Gaussian sample with known mean*

$\underline{X} = [X_1, \ldots, X_n]^T$ i.i.d., $X_1 \sim \mathcal{N}(\mu, \sigma^2)$, μ known.

Again three cases of interest:

$$\underbrace{\begin{array}{l} H_0 : \sigma^2 = \sigma_o^2 \\ H_1 : \sigma^2 > \sigma_o^2 \end{array}}_{\text{Case I}} \quad \underbrace{\begin{array}{l} H_0 : \sigma^2 = \sigma_o^2 \\ H_1 : \sigma^2 < \sigma_o^2 \end{array}}_{\text{Case II}} \quad \underbrace{\begin{array}{l} H_0 : \sigma^2 = \sigma_o^2 \\ H_1 : \sigma^2 \neq \sigma_o^2 \end{array}}_{\text{Case III}}$$

Solution

STEP 1: find MP-LRT for fixed σ^2

Approach 1: work problem directly from entire random data sample $\underline{X}$.

The likelihood ratio depends on σ^2 and, for fixed value of σ^2, is given by:

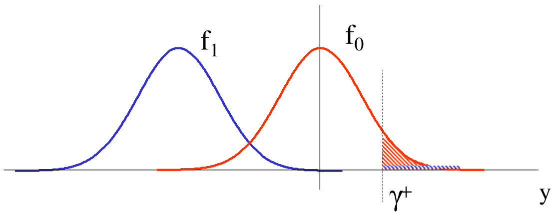


Figure 101: *The single sided MP-LRT for* $H_0 : \mu = 0$ *vs.* $H_1 : \mu > 0$ *fails to detect negative signal.*

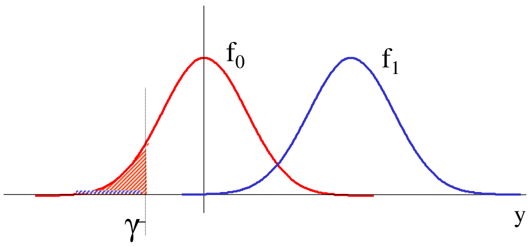


Figure 102: *The single sided MP-LRT for* $H_0 : \mu = 0$ *vs.* $H_1 : \mu > 0$ *fails to detect positive signal.*

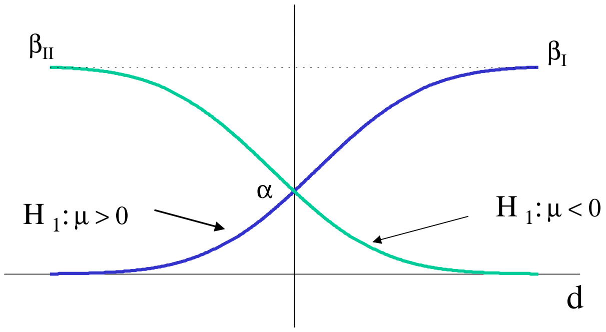


Figure 103: *The power curve of Case I or Case II MP-LRT's for double sided hypotheses is biased over range $-\infty < d < \infty$.*

$$\begin{aligned}
\Lambda(\sigma^2) &= \frac{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^n (X_k-\mu)^2\right)}{\left(\frac{1}{\sqrt{2\pi\sigma_o^2}}\right)^n \exp\left(-\frac{1}{2\sigma_o^2}\sum_{k=1}^n (X_k-\mu)^2\right)} \\
&= \left(\frac{\sigma_o^2}{\sigma^2}\right)^{n/2} \exp\left(\frac{\sigma^2-\sigma_o^2}{2\sigma^2\sigma_o^2}\sum_{k=1}^n (X_k-\mu)^2\right) \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta
\end{aligned}$$

Which is monotone increasing in the quantity

$$(\sigma^2-\sigma_o^2)\sum_{k=1}^n (X_k-\mu)^2$$

Thus we obtain MP-LRT

$$(\sigma^2-\sigma_o^2)\,\frac{n\,\hat{\sigma}_\mu^2}{\sigma_o^2} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

where

$$\hat{\sigma}_\mu^2 = n^{-1}\sum_{k=1}^n (X_k-\mu)^2$$

Approach 2: work problem based on sufficient statistic for σ^2

$$T(\underline{X}) = \sum_{i=1}^{n}(X_i - \mu)^2$$

The density of $T(\underline{X})/\sigma^2$ is Chi-square with n d.f.
$\Rightarrow$ using standard transformation of variables formula, p.d.f. of T is:

$$\begin{aligned} f(T;\sigma^2) &= \sigma^{-2}\, f_\chi(T\,\sigma^{-2}) \\ &= \sigma^{-2}\,\frac{1}{2^{n/2}\Gamma(n/2)}\, e^{-T/(2\sigma^2)}(T/\sigma^2)^{n/2-1} \end{aligned}$$

Hence MP-LRT is

$$\Lambda(\sigma^2) = \left(\frac{\sigma_o^2}{\sigma^2}\right)^{n/2} \exp\left\{\frac{\sigma^2-\sigma_o^2}{2\sigma^2\sigma_o^2}\, T\right\} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

Which is monotone in $(\sigma^2 - \sigma_o^2)\, T$.
Thus we obtain MP-LRT

$$(\sigma^2 - \sigma_o^2)\,\frac{n\,\hat{\sigma}_\mu^2}{\sigma_o^2} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

where again

$$\hat{\sigma}_\mu^2 = n^{-1}\sum_{k=1}^{n}(X_k - \mu)^2$$

CASE I: Single sided alternative $H_1 : \sigma^2 > \sigma_o^2$
In this case MP-LRT is simply

$$T(\underline{X}) = \frac{n\,\hat{\sigma}_\mu^2}{\sigma_o^2} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma^+$$

or equivalently, we have a square law detector

$$T(\underline{X}) = \frac{1}{\sigma_o^2}\sum_{i=1}^{n}(X_i - \mu)^2 \underset{H_0}{\overset{H_1}{\gtrless}} \gamma^+$$

Under H_0, $X_k \sim \mathcal{N}(\mu, \sigma_o^2)$ so the test statistic $T(\underline{X})$ is Chi-square with n d.f.
Therefore

$$\alpha = P_0(T(\underline{X}) > \gamma^+) = 1 - \chi_n(\gamma^+)$$

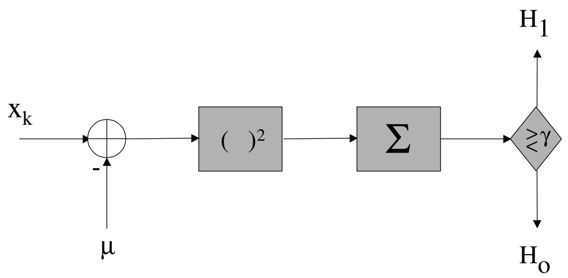


Figure 104: *Optimal detector for increase in Gaussian variance is a memoryless non-linear device (squarer) followed by a summer and decision mechanism. This detector has been called the square law detector and the energy detector.*

and

$$\gamma^+ = \chi_n^{-1}(1-\alpha)$$

Hence MP-LRT is

$$\frac{n\,\hat{\sigma}_\mu^2}{\sigma_o^2} \overset{H_1}{\underset{H_0}{\gtrless}} \chi_n^{-1}(1-\alpha)$$

which is UMP against any $\sigma^2 > \sigma_o^2$ for known μ.

Power: since $\frac{\sigma_o^2}{\sigma^2}\,\frac{n\hat{\sigma}_\mu^2}{\sigma_o^2} = \chi_n$

$$\beta = P_1\left(n\,\hat{\sigma}_\mu^2/\sigma_o^2 > \gamma^+\right) = 1 - \chi_n\left(\frac{\sigma_o^2}{\sigma^2}\,\chi_n^{-1}(1-\alpha)\right)$$

CASE II: Single sided alternative $H_1 : \sigma^2 < \sigma_o^2$

Find that MP-LRT has form

$$\frac{n\,\hat{\sigma}_\mu^2}{\sigma_o^2} \overset{H_0}{\underset{H_1}{\gtrless}} \gamma^-$$

where now

$$\gamma^- = \chi_n^{-1}(\alpha)$$

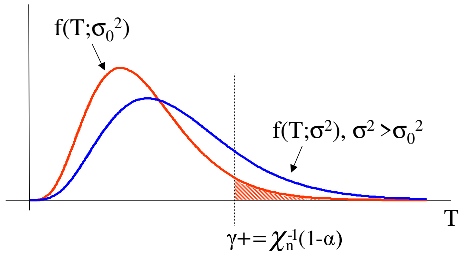


Figure 105: *Density functions under H_0 and H_1 of optimal UMP test statistic for testing against $\sigma^2 > \sigma_o^2$ for known mean μ. Threshold γ^+ is determined by the $1-\alpha$ quantile of the H_0 density.*

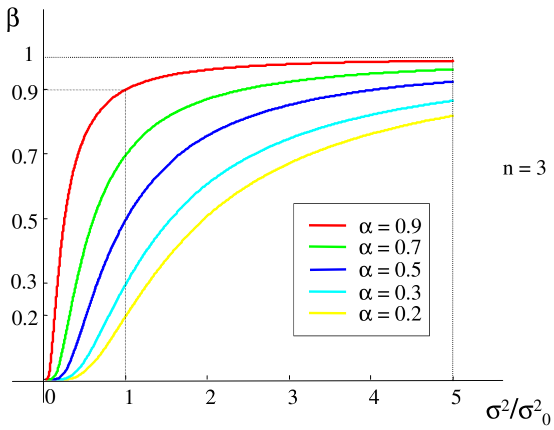


Figure 106: *Power curves for one sided test of variance $\sigma^2 > \sigma_o^2$ for known mean μ with i.i.d. Gaussian observations for various values of σ^2/σ_o^2 and $n = 3$.*

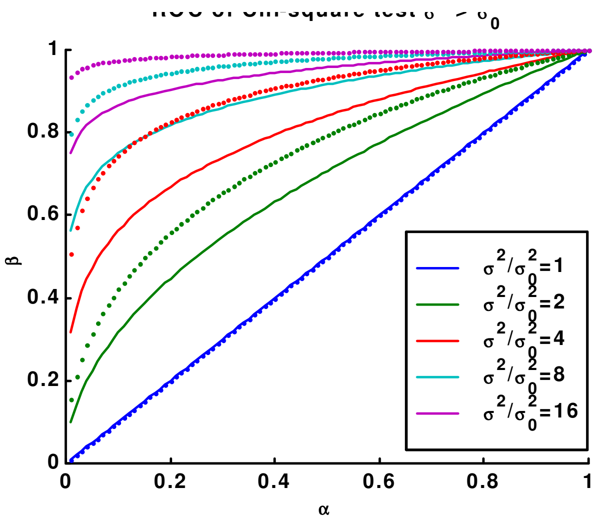


Figure 107: *ROC curves for one sided test of variance $\sigma^2 > \sigma_o^2$ with i.i.d. Gaussian observations for various values of σ^2/σ_o^2 and $n = 3$.*

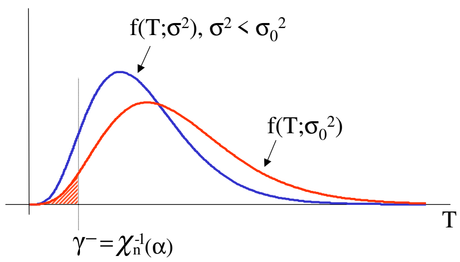


Figure 108: *Density functions under H_0 and H_1 of optimal UMP test statistic for testing against $\sigma^2 < \sigma_o^2$ for known mean μ. Threshold γ^+ is determined by the $1-\alpha$ quantile of the H_0 density.*

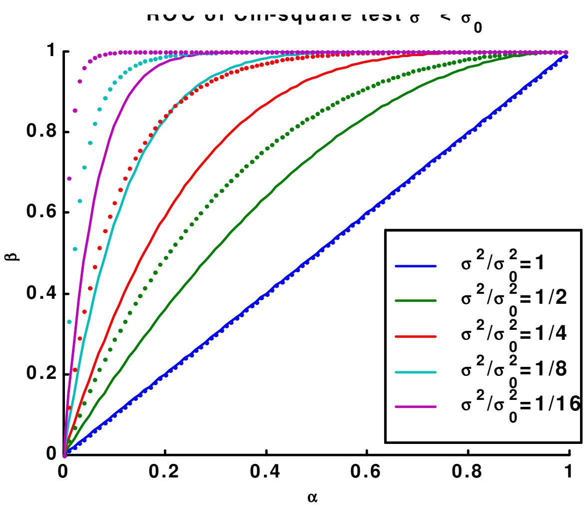


Figure 109: *ROC curves for one sided test of variance $\sigma^2 < \sigma_o^2$ with i.i.d. Gaussian observations for various values of σ^2/σ_o^2 and $n = 3$.*

So that we have UMP test against $\sigma^2 < \sigma_o^2$ for known μ

Power:

$$\beta = 1 - \chi_n\left(\frac{\sigma_o^2}{\sigma^2}\,\chi_n^{-1}(\alpha)\right)$$

Case III: Double sided alternative $H_1 : \sigma^2 \neq \sigma_o^2$

No UMP exists.

Example 39 *One sided test on median of Cauchy density*

Assume $X_1, \ldots, X_n$ i.i.d. with marginal density

$$f(x_1; \theta) = \frac{1}{\pi}\frac{1}{1 + (x_1 - \theta)^2}$$

Objective: investigate existence of UMP test for

$$\begin{aligned} H_0 &: \theta = 0 \\ H_1 &: \theta > 0 \end{aligned}$$

Step 1: First find LRT for fixed $\theta > 0$

$$\Lambda(\theta) \quad = \quad \frac{f(x;\theta)}{f(x;0)} = \prod_{i=1}^{n} \frac{1 + x_i^2}{1 + (x_i - \theta)^2} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

where we have explicitly indicated the dependency of Λ on θ in the notation (dependence of Λ on $\underline{x}$ is suppressed). For the special case of a single sample ($n = 1$):

$$\Lambda(\theta) = \frac{1 + x_1^2}{1 + (x_1 - \theta)^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

Step 2: The decision region depends on θ even if $\theta > 0$ (See exercises). Therefore, in this case no UMP exists even for the one sided hypothesis!

8.2 GENERAL CONDITION FOR UMP TESTS: MONOTONE LIKELIHOOD RATIO

Consider testing a composite alternative hypothesis against a simple null hypothesis

$$H_0 : \theta = \theta_0 \tag{107}$$
$$H_1 : \theta \in \Theta_1, \tag{108}$$

Then if the likelihood ratio is monotone there exists a UMP test. We specialize to a 1D parameter $\theta \in \mathbb{R}$. Let $f(x; \theta)$ have Fisher Factorization

$$f(x; \theta) = g(T, \theta)h(x),$$

where T is a sufficient statistic. The Carlin-Rubin monotonicity theorem states that [39]

Monotone Likelihood Ratio Theorem: an UMP test of (108) at any level $\alpha \in [0, \alpha]$ exists if the likelihood ratio is either monotone increasing or monotone decreasing in T for all $\theta \in \Theta_1$

$$\Lambda = \frac{f(x; \theta)}{f(x; \theta_0)} = \frac{g(T; \theta)}{g(T; \theta_0)} = \Lambda_\theta(T)$$

To prove the theorem note that the MP test for a simple alternative $H_1 : \theta = \theta_1$ is

$$\Lambda_{\theta_1}(T) \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

which is equivalent to a test that compares T to a threshold $\gamma = \Lambda_{\theta_1}^{-1}(\eta)$

$$T \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \quad \text{(increasing } \Lambda\text{)}$$

$$T \underset{H_1}{\overset{H_0}{\gtrless}} \gamma \quad \text{(decreasing } \Lambda\text{)}$$

For the special case of a one sided alternative H_1:

$$H_1 : \theta > \theta_0, \quad \text{or} \quad H_1 : \theta < \theta_0,$$

there are several densities that satisfy the monotone LR condition and which therefore admit UMP tests, which can be obtained by selecting a $\theta_1 > \theta_0$ and deriving the MP-LRT of H_0 versus the simple alternative $H_1 : \theta = \theta_1$. The following are examples

1. x i.i.d. sample from 1D exponential family,
2. In particular: Gaussian, Bernoulli, Exponential, Poisson, Gamma, Beta
3. x i.i.d. sample from a Uniform density $\mathcal{U}(0, \theta)$
4. x i.i.d. sample from noncentral-t, noncentral Fisher F
5. x i.i.d. sample from shifted Laplace, logistic

In fact, it can be shown [39] that the monotone LR condition guarantees that the MP-LRT is UMP wrt $H_o : \theta < \theta_0$ too!

There are lots of situations where the monotone LR does not hold. For example, the following

1. Gaussian density with single sided H_1 on mean but having unknown variance
2. Cauchy density with single sided H_1
3. Exponential family with double sided H_1

8.3 COMPOSITE HYPOTHESIS DETECTION STRATEGIES

Here it is desired to test doubly composite

$$\begin{aligned} H_0 : \theta &\in \Theta_0 \\ H_1 : \theta &\in \Theta_1 \end{aligned}$$

Now, most fixed detectors will have both P_F and P_D varying as functions of $\theta \in \Theta_0$ and $\theta \in \Theta_1$, respectively.

Recall that for composite H_0 we say that test ϕ is of level α if

$$\max_{\theta_0 \in \Theta_0} P_F(\theta_0) \leq \alpha$$

where $P_F(\theta) = E_{\theta_0}[\phi]$

Two classes of strategies:

1. Optimize alternative detection criterion
2. Constrain form of detector to a class for which UMP may exist

8.4 MINIMAX TESTS

A conservative approach to testing composite hypotheses would be to maximize worst case power under a constraint on worst case false alarm. This approach is called a minimax strategy and leads to conservative but minimax optimal tests. Minimax approaches are not very widespread in signal processing applications due to their overly conservative performance and their often difficult implementation.

Objective: find level α test which satisfies the constraint:

$$\max_{\theta \in \Theta_0} E_\theta[\phi] \leq \alpha,$$

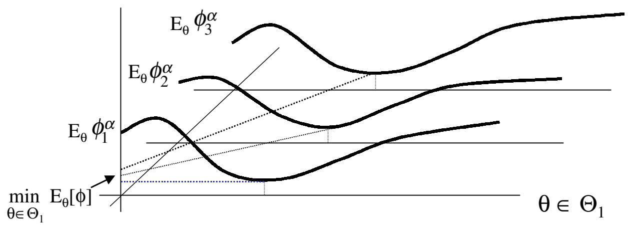


Figure 110: *Various power curves for different test functions and their minima over the unknown parameter θ varying over H_1 parameter space Θ_1. Minimax NP test ϕ_3^α maximizes minimum power.*

and maximizes the worst case power

$$\min_{\theta\in\Theta_1} E_\theta[\phi].$$

METHOD OF SOLUTION: find "least favorable" densities

Simplifying assumption: Θ discrete parameter space.

Fundamental identity on the mean [22]: For any summable sequence $\{a(k)\}_k$ and any probability distribution $\{p(k)\}_k$

$$\min_k a(k) \leq \sum_k a(k)p(k) \leq \max_k a(k)$$

with equality when $p(k)$ = delta function concentrated on $\text{argmin}_k a(k)$ and $\text{argmax}_k a(k)$. Therefore

$$\min_k a(k) = \min_{\{p(k)\}} \sum_k a(k)p(k), \quad \max_k a(k) = \max_{\{p(k)\}} \sum_k a(k)p(k)$$

* Let $\{p_0(\theta)\}$ be an arbitrary probability distribution on Θ_0

* Let $\{p_1(\theta)\}$ be an arbitrary probability distribution on Θ_1

Then the worst case $P_F(\theta)$ and $P_M(\theta)$ can be expressed as worst case average P_F and P_M

$$\max_{\theta\in\Theta_0} E_\theta[\phi] = \max_{p_0} \sum_{\theta\in\Theta_0} E_\theta[\phi]p_0(\theta) = \sum_{\theta\in\Theta_0} E_\theta[\phi]p_0^*(\theta)$$

$$\min_{\theta\in\Theta_1} E_\theta[\phi] = \min_{p_1} \sum_{\theta\in\Theta_1} E_\theta[\phi]p_1(\theta) = \sum_{\theta\in\Theta_1} E_\theta[\phi]p_1^*(\theta)$$

where

* p_0^* maximizes the false alarm probability

* p_1^* minimizes the detection probability (power)

Define "least favorable pair" of densities

$$
\begin{aligned}
f_0^*(x) &= \sum_{\theta\in\Theta_0} f(x;\theta)p_0^*(\theta) \\
f_1^*(x) &= \sum_{\theta\in\Theta_1} f(x;\theta)p_1^*(\theta)
\end{aligned}
$$

Then the minimax objective reduces to

Constraint:

$$E_0^*[\phi] = \int_{\mathcal{X}} \phi(x)\ f_0^*(x)dx \leq \alpha$$

Maximize:

$$E_1^*[\phi] = \int_{\mathcal{X}} \phi(x)\ f_1^*(x)dx$$

Which, corresponds to finding a MP test of level α for the derived simple hypotheses

$$
\begin{aligned}
H_0^* : X &\sim f_0^* \\
H_1^* : X &\sim f_1^*
\end{aligned}
$$

Hence minimax NP test is the LRT

$$\frac{f_1^*(x)}{f_0^*(x)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \eta$$

where threshold η is chosen to satisfy:

$$\int_{\mathcal{X}} \phi^*(x) f_0^*(x)dx = \alpha$$

Observations

* Minimax NP test is an optimal Bayes test for random θ over Θ_1 and Θ_0 but without prior probabilities on H_0 and H_1.

* Performance of minimax NP test can be overly conservative, especially if least favorable priors concentrate on atypical values of θ.

* Least favorable priors p_1^*, p_0^* may be difficult to find in practice

$\Rightarrow$ Helpful facts concerning f_1^*, f_0^* [16]:

* p_0^* and p_1^* make H_0^* and H_1^* the most difficult to discriminate

* p_1^* and p_0^* can each assume at most two values over Θ_1 and Θ_0

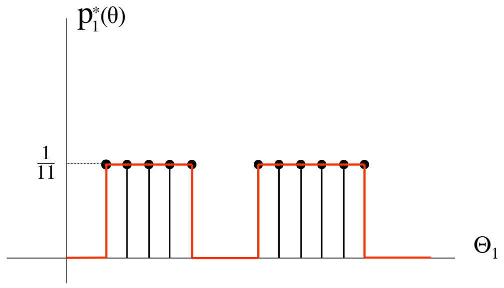


Figure 111: *Least favorable density* $p_1^*(\theta)$ *is piecewise constant over* Θ_1.

Specifically, there exists a subset Θ_0^+ of Θ_0 such that

$$p_0^*(\theta) = \begin{cases} q, & \theta \in \Theta_0^+ \\ 0, & \theta \in \Theta_0 - \Theta_0^+ \end{cases}$$

where q is equal to the volume of Θ_0^+

$$q = \begin{cases} \int_{\Theta_0^+} d\theta, & \Theta_0 \text{ cts.} \\ \sum_{\theta \in \Theta_0^+}, & \Theta_0 \text{ discrete} \end{cases}$$

and similarly for p_1^*.

Examples of minimax tests will be explored in the exercises.

8.5 LOCALLY MOST POWERFUL (LMP) SINGLE SIDED TEST

Main idea: if we can't find a UMP over the entire set $\theta > \theta_0$ then perhaps we can find a test that remains MP over small perturbations, e.g., $\theta \in (\theta_0, \theta_0 + \Delta]$ with $(0 < \Delta \ll 1)$, from H_0 . First we consider single sided case and 1D parameter θ

$$\begin{aligned} H_0 : \theta &= \theta_0 \\ H_1 : \theta &> \theta_0 \end{aligned}$$

The idea is simple. Referring to Fig. 112, we recall that the power curve of a good test increases as a function of θ. Therefore, it makes sense to try and find a test that will maximize the rate of increase near θ_0. This leads to the definition:

Definition 2 *A locally most powerful (LMP) test* ϕ *of level* α *has power curve that maximizes slope of* $\beta(\theta)$ *at* $\theta = \theta_0$

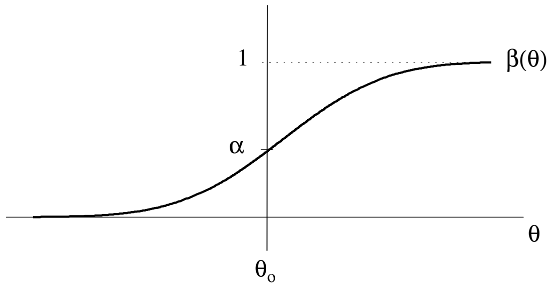


Figure 112: *Typical power curve $\beta(\theta)$. LMP test of $H_1 : \theta > \theta_o$ seeks to maximize the slope of the power curve at point $\theta = \theta_o$.*

We can formulate the LMP testing strategy ϕ by posing it as the following optimization:

Constrain: $E_{\theta_0}[\phi] \leq \alpha$

Maximize: $\frac{d}{d\theta_0} E_{\theta_0}[\phi]$

Similarly to the derivation of NPL in the previous chapter, we obtain the solution ϕ to this optimization as the test

$$\phi_{LMP}(x) = \begin{cases} 1, & df(x;\theta_0)/d\theta_0 > \eta f(x;\theta_0) \\ q, & df(x;\theta_0)/d\theta_0 = \eta f(x;\theta_0) \\ 0, & df(x;\theta_0)/d\theta_0 < \eta f(x;\theta_0) \end{cases}$$

Or for short

$$\Lambda_{\rm LMP}(x) = \frac{df(x;\theta_0)/d\theta_0}{f(x;\theta_0)} \overset{H_1}{\underset{H_0}{\gtrless}} \eta$$

where η is selected to satisfy constraint (possibly with randomization)

$$E_{\theta_0}[\phi] \leq \alpha$$

To prove this is quite simple if we follow the Lagrange multiplier approach that was used to derive the MP test of Lemma 1. First, note that we can express $\frac{d}{d\theta_0} E_{\theta_0}[\phi]$ using the "Girsanov representation" and a relation for the derivative of the logarithm function

$$\begin{aligned} \frac{d}{d\theta_0} E_{\theta_0}[\phi] &= \frac{d}{d\theta_0} \int \phi(x) f_{\theta_0}(x) dx \\ &= \frac{d}{d\theta_0} \int \phi(x) f_{\theta_0}(x) dx \end{aligned}$$

$$
\begin{aligned}
&= \int \phi(x) \frac{d}{d\theta_0} f_{\theta_0}(x) dx \\
&= \int \phi(x) \left(\frac{d}{d\theta_0} \ln f_{\theta_0}(x) \right) f_{\theta_0}(x) dx \\
&= E_{\theta_0}\left[\phi \left(\frac{d}{d\theta_0} \ln f_{\theta_0} \right) \right].
\end{aligned}
$$

Therefore, the Lagrangian associated with our constrained maximization problem is simply written as:

$$
\frac{d}{d\theta_0} E_{\theta_0}[\phi] + \eta(\alpha - E_{\theta_0}[\phi]) = E_{\theta_0}\left[\phi \left(\frac{d}{d\theta_0} \ln f_{\theta_0} - \eta \right) \right] + \eta\alpha,
$$

which is obviously maximized by selecting $\phi = \phi_{LMP}$ given above.

There is a close connection between the LMP and maximum likelihood estimation. Assuming that we have set $\eta = 0$ we can write the LMP test in an equivalent form

$$
\Lambda_{\text{LMP}} = \frac{d}{d\theta_o} \ln f(x; \theta_0) \overset{H_1}{\underset{H_0}{\gtrless}} 0.
$$

Thus we decide H_1 if the slope of the likelihood function is positive at $\theta = \theta_0$. Such a situation occurs when the log-likelihood function is strictly concave and the MLE $\hat{\theta}$ is greater than θ_0, i.e. the MLE provides good evidence that H_1 is true! If $\eta > 0$ then the slope at θ_o has to be both large and positive, providing even stronger evidence that $\theta > \theta_0$.

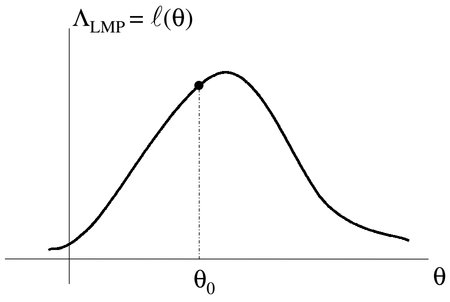


Figure 113: *LMP test $H_1 : \theta > \theta_o$ decides H_0 if θ_0 is near stationary point of log likelihood function $l(\theta)$.*

Example 40 *Gaussian one sided test against zero mean*

Find: differential LR has the form

$$\begin{aligned} \frac{df(x;\theta)/d\theta}{f(x;\theta)} &= \frac{d}{d\theta}\ln f(x;\theta) \\ &= \frac{\sum_{i=1}^n (X_i-\theta)}{\sigma^2} \end{aligned}$$

LMP for testing $\theta = \theta_0 = 0$ vs. $\theta > 0$ is therefore:

$$\sum_{i=1}^n X_i \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

or level α LMP is the linear UMP test obtained before

$$\frac{\sqrt{n}\,\overline{X_i}}{\sigma} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

Example 41 *Cauchy one sided test against zero median (ctd)*

Find: differential LR has the form

$$\frac{df(x;\theta)/d\theta}{f(x;\theta)} = 2\sum_{i=1}^n \frac{X_i-\theta}{1+(X_i-\theta)^2}$$

For $\theta = \theta_0 = 0$ LMP test is therefore:

$$T(\underline{X}) = \sum_{i=1}^n \frac{X_i}{1+X_i^2} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

Test statistic $T(\underline{X})$ is sum of i.i.d. r.v.s with mean 0 and variance 1/8 under H_0. Therefore threshold γ can be found via CLT for large n:

$$\gamma = \sqrt{n/8}\,\mathcal{N}^{-1}(1-\alpha)$$

Example 42 *Testing for positive mean of Laplace distribution*

* $\underline{X} = [X_1, \ldots, X_n]$ i.i.d,

$$X_i \sim f(x;\theta) = \frac{a}{2}e^{-a|x-\theta|}, \quad a > 0$$

Log-likelihood function takes the form:

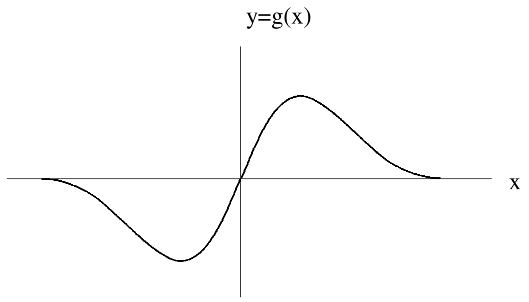


Figure 114: *Memoryless non-linearity* $g(x) = x/(1+x^2)$ *input-output characteristic for LMP test of one sided test against zero median for a Cauchy r.v.*

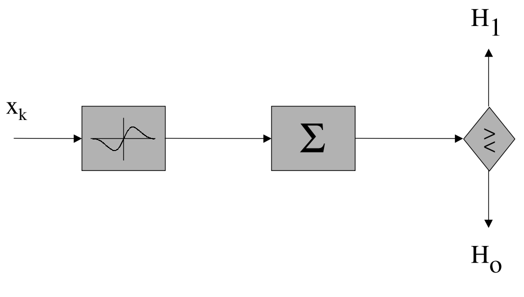


Figure 115: *Optimal detector for positive Cauchy median is a memoryless non-linearity followed by a summer and decision mechanism.*

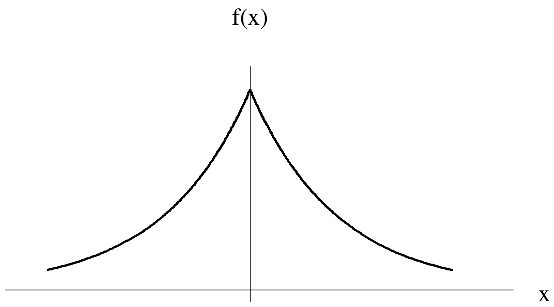


Figure 116: *Laplace density* $f(x) = ae^{-a|x-\theta|}/2$. *Width, as measured by where* $f(x)$ *falls to* $1/e$ *of its peak, is* $2/a$.

$$\begin{aligned}
\ln f(\underline{x};\theta) &= -a\sum_{i=1}^{n}|X_i-\theta| + n\ln\frac{a}{2} \\
&= -a\sum_{X_i>\theta}(X_i-\theta) + a\sum_{X_i<\theta}(X_i-\theta) + c \\
&= a\theta\,(n_+ - n_-) + b(\theta)
\end{aligned}$$

where

$$n_+ = \#\ X_i > \theta, \qquad n_- = \#\ X_i < \theta = n - n_+$$

Note: $b(\theta)$ is piecewise constant function

Find: differential LR has the form

$$\frac{df(x;\theta_o)/d\theta_o}{f(x;\theta_o)} = a(n_+ - n_-)$$

LMP is therefore:

$$T(\underline{X}) = n_+ - n_- \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

or equivalently, in a form more comparable to the Cauchy and Gaussian examples (41) and (37) having $\theta_o = 0$:

$$T(\underline{X}) = \sum_{i=1}^{n} \mathrm{sgn}(X_i) \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

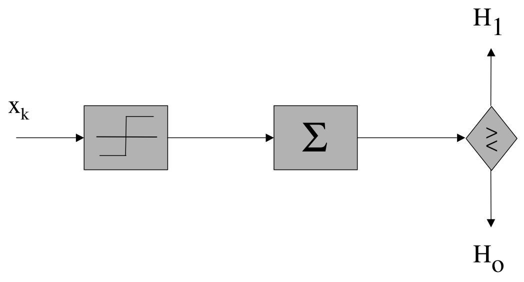


Figure 117: *LMP detector for testing positive mean $\theta > 0$ for a Laplace r.v. is composed of a summer and memoryless non-linearity.*

PERFORMANCE:

$T(\underline{X})$ is a discrete shifted Binomial r.v.

$$T(\underline{X}) = \sum_{i=1}^{n} (2b_i - 1) = 2B(n, p) - n$$

where b_i are i.i.d. Bernoulli r.v.'s with parameter

$$p = P_\theta(b_i = 1) = P_\theta(X_i > 0)$$

$\Rightarrow$ Randomized test is necessary to set false alarm.

$$\alpha = P_0(T(\underline{X}) > \gamma_-) + q(\alpha_+ - \alpha_-)$$

where α_- and γ_- are related by

$$\alpha_- = P_0(\underbrace{T(\underline{X})}_{2B(n,\frac{1}{2})-n} > \gamma_-) = 1 - B_{n,p}\left(\frac{\gamma_- + n}{2}\right)$$

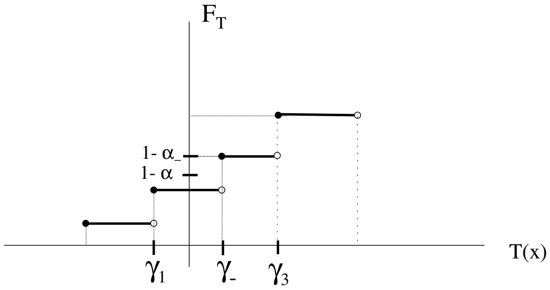


Figure 118: *The CDF of test statistic is staircase function (value of F_T over $\gamma_1 \leq T(\underline{X}) < \gamma_-$ is $1 - \alpha_+$). Randomization is necessary for meeting FA constraint.*

and the randomization parameter q is as usual

$$q = \frac{\alpha - \alpha_-}{\alpha_+ - \alpha_-}$$

8.6 MOST POWERFUL UNBIASED (MPU) TESTS

Recall: a test ϕ of level α is an unbiased test if

$$E_\theta[\phi] \geq \alpha, \quad \text{all } \theta \in \Theta_1.$$

A test ϕ of level α is uniformly MPU (UMPU) if for all $\theta \in \Theta_1$ its power function dominates that of all other unbiased tests of level α. By restricting the class of competing tests there is hope that a MP test may emerge among them. Unfortunately this is not much more frequent than in the unrestricted case. For more details on the theory and practice of unbiased testing see Lehmann [39].

8.7 LOCALLY MOST POWERFUL UNBIASED DOUBLE SIDED TEST

Consider double sided hypotheses:

$$\begin{aligned} H_0 : \theta &= \theta_0 \\ H_1 : \theta &\neq \theta_0 \end{aligned}$$

Observe: The power function of a good unbiased level α test ϕ should have global minimum at $\theta = \theta_0$.

Locally unbiased test optimization for 1D parameter θ

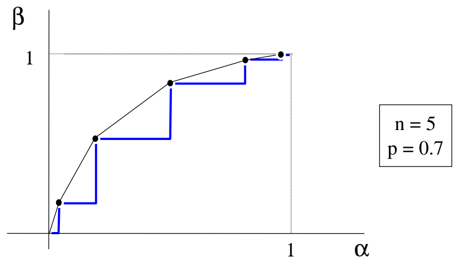


Figure 119: *ROC curve of LMP detector for testing positive mean $\theta > 0$ for a Laplace r.v.*

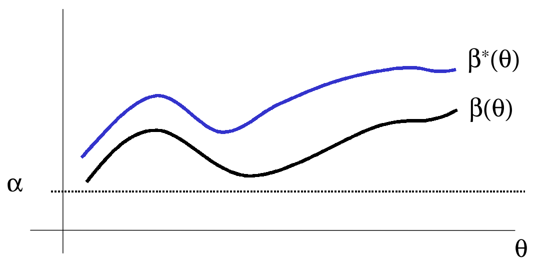


Figure 120: *Power curve of most powerful unbiased test (MPU) dominates that of all other unbiased tests of the same FA level.*

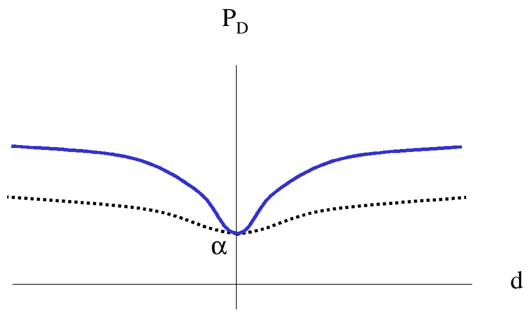


Figure 121: *Power curve of a good locally unbiased test has minimum at α with maximum curvature.*

Constraints:

$$E_{\theta_0}[\phi] \leq \alpha, \qquad \frac{d}{d\theta_0} E_{\theta_0}[\phi] = 0. \tag{109}$$

Subject to these constraints want to maximize curvature at θ_0

$$\frac{d^2}{d\theta_0^2} E_{\theta_0}[\phi].$$

Using Lagrange multipliers it is easily shown that the test function ϕ which solves this constrained maximization problem has the form:

$$\phi(x) = \left\{ \begin{array}{ll} 1, & d^2 f(x;\theta_0)/d\theta_0^2 > \eta(f(x;\theta_0) + \rho \, df(x;\theta_0)/d\theta_0) \\ q & d^2 f(x;\theta_0)/d\theta_0^2 = \eta(f(x;\theta_0) + \rho \, df(x;\theta_0)/d\theta_0) \\ 0 & d^2 f(x;\theta_0)/d\theta_0^2 < \eta(f(x;\theta_0) + \rho \, df(x;\theta_0)/d\theta_0) \end{array} \right. \tag{110}$$

where ρ, η, q are selected to satisfy the two constraints.

In some cases, one can meet the constraints by selecting $\rho = 0$ and varying only $q \in [0,1]$ and $\eta \in [0, \infty)$. In this situation, the locally optimal test (110) reduces to the simpler (randomized) LRT form

$$\frac{d^2 f(x;\theta_0)/d\theta_0^2}{f(x;\theta_0)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \eta.$$

Example 43 *Double sided test against zero mean of Gaussian sample with known variance*

Step 1: Find derivatives of p.d.f. of sufficient statistic $\overline{X}$ (Here for clarity we define its pdf as $f_{\overline{X}}(v;\mu)$ for $v \in \mathbb{R}$).

$$\begin{aligned}
f_{\overline{X}}(v;\mu) &= \frac{1}{\sqrt{2\pi\sigma^2/n}} e^{-\frac{(v-\mu)^2}{2\sigma^2/n}} \\
df_{\overline{X}}(v;\mu)/d\mu &= \left(n/\sigma^2\right)\ (v-\mu) f_{\overline{X}}(v;\mu) \\
d^2 f_{\overline{X}}(v;\mu) d\mu^2 &= \left(n/\sigma^2\right)\ [n/\sigma^2 (v-\mu)^2 - 1] f_{\overline{X}}(v;\mu)
\end{aligned}$$

Thus LMPU LRT is

$$\frac{\overline{X}^2 - \sigma^2/n}{\sigma^2/n + \rho\overline{X}} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

Step 2: Select ρ, η to satisfy constraints

First we attempt to satisfy constraints with $\rho = 0$ and η a free variable.

For this case LMPU LRT reduces to

$$|\overline{X}| \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma \tag{111}$$

Since

$$\overline{X} \sim \mathcal{N}(0, \sigma^2/n), \qquad \text{under } H_0$$

we have

$$\alpha = 1 - P_0(-\gamma < \overline{X} \le \gamma) = 2(1 - \mathcal{N}(\gamma\sqrt{n}/\sigma))$$

Or

$$\gamma = \frac{\sigma}{\sqrt{n}} \mathcal{N}^{-1}(1 - \alpha/2)$$

Success! We can set threshold γ to achieve arbitrary $P_F = \alpha$ with $\rho = 0$ and without randomization. Of course it still must be verified that the test (111) satisfies the second constraint in (109) which is that $d/d\mu P_D(\mu)|_{\mu=0} = 0$. This can be shown by establishing symmetry of the power function about $\mu = 0$. The details are left as an exercise.

Equivalent form of locally-unbiased test

$$\frac{1}{n} \left| \sum_{i=1}^n X_i \right| \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

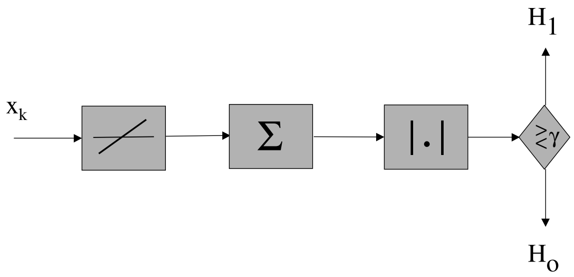


Figure 122: *Locally best unbiased double-sided test for non-zero Gaussian mean is a memoryless non-linearity followed by a summer and decision device.*

Power:

Since

$$\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n), \quad \text{under } H_1$$

$$\begin{aligned} P_D &= 1 - P_\mu(-\gamma < \overline{X} \le \gamma) \\ &= 1 - [\mathcal{N}(\sqrt{n}\,(\gamma-\mu)/\sigma) - \mathcal{N}(\sqrt{n}\,(-\gamma-\mu)/\sigma)] \end{aligned}$$

$$= 1 - \mathcal{N}(\mathcal{N}^{-1}(1-\alpha/2) - d) - \mathcal{N}(\mathcal{N}^{-1}(1-\alpha/2) + d)$$

where as usual:

* $d = \sqrt{n}\,\mu/\sigma$ is detectability index

Remark:

* It can be shown that in the Gaussian example above the LMPU test is actually UMPU. See Ferguson [16].

The LMPU strategy can in principle be extended to multiple parameters as follows. Assume

$$\underline{\theta} = [\theta_1, \ldots, \theta_p]^T$$

and let's test the hypotheses:

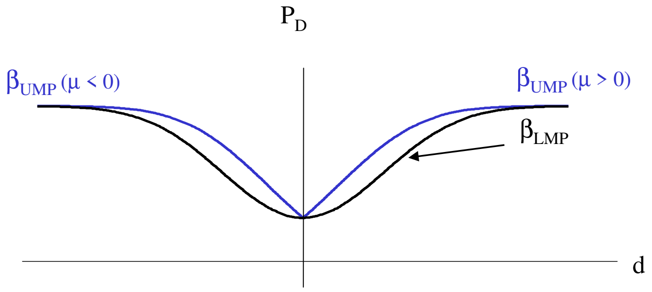


Figure 123: *Power curve of LMPU test for for non-zero Gaussian mean with known variance as a function of values of d.*

$$H_0 : \underline{\theta} = \underline{\theta}_0$$
$$H_1 : \underline{\theta} \neq \underline{\theta}_0$$

Constraints:

$$E_{\underline{\theta}_0}[\phi] \leq \alpha, \quad \nabla_{\underline{\theta}_0} E_{\underline{\theta}_0}[\phi] = \underline{0}) \tag{112}$$

Maximize:

$$\mathrm{trace}\left\{\nabla^2_{\underline{\theta}_0} E_{\underline{\theta}_0}[\phi]\right\}$$

where trace $\{\mathbf{A}\}$ denotes trace of matrix $\mathbf{A}$.

This is a similar optimization as we encountered in proving the Neyman Pearson Lemma. However, now there are $p+1$ constraints as indicated in (112). One of them is due to the false alarm constraint and p of them are due to constraining the gradient vector to zero. The optimal test can be found by applying Lagrange multipliers and has the form

$$\frac{\mathrm{trace}\left\{\nabla^2_{\underline{\theta}_0} f(x;\underline{\theta}_0)\right\}}{f(x;\underline{\theta}_0) + \sum_{i=1}^p \rho_i \partial f(x;\underline{\theta}_0)/\partial\theta_{0i}} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \eta,$$

where $\rho_1, \ldots, \rho_p, \eta$ are selected to satisfy the constraints (possibly with randomization).

8.8 CFAR DETECTION

A sometimes reasonable condition is to require that tests have constant false alarm rate (CFAR), i.e. constant $P_F(\theta)$ over $\theta \in \Theta_0$. Then one attemps to find a UMP CFAR test. The setup is as follows:

Constraint:

$$E_\theta[\phi] = \alpha, \quad \theta \in \Theta_0$$

Maximize:

$$E_\theta[\phi], \quad \theta \in \Theta_1$$

A effective methodology for finding CFAR tests is by the use of invariance principles [33]. CFAR tests are also known as similar tests and for more information see [39].

8.9 INVARIANT TESTS

Consider the general case where we have partition of $\underline{\theta}$

$$\underline{\theta} = [\varphi_1, \ldots, \varphi_p, \quad \underbrace{\xi_1, \ldots, \xi_q}_{\text{nuisance parameters}} \quad]^T$$

and $X \sim f(x; \underline{\varphi}, \underline{\xi})$

It is desired to test single sided hypotheses

$$\begin{aligned} H_0 &: \underline{\varphi} = 0, \quad \underline{\xi} = \underline{\xi}_o \\ H_1 &: \underline{\varphi} > 0, \quad \underline{\xi} = \underline{\xi}_o \end{aligned}$$

where $\underline{\xi}_o \in \mathbb{R}^q$ is *unknown* $\Rightarrow$ UMP does not usually exist.

Invariant tests seek to find a transformation (compression) of the data

$$Z = Z(X)$$

which satisfies:

Property 1. Z contains (almost) as much information concerning φ as X

Property 2. Distribution of Z is not a function of $\underline{\xi}$.

Due to Property 2, if we throw away x and retain only

$Z \sim f(z; \varphi)$

then we are back to testing simpler hypotheses for which an UMP may exist

$$\begin{aligned} H_0 &: \underline{\varphi} = 0 \\ H_1 &: \underline{\varphi} > 0 \end{aligned}$$

The theory of optimal invariant tests is treated in detail in [33] in which invariance is referred to as exact robustness.

For now we concentrate on a particular suboptimal "invariant" approach

8.10 GENERALIZED LIKELIHOOD RATIO TEST

We now turn to one of the most prevalent methods of dealing with detection for composite hypotheses. Unlike the previous methods, which were all motivated by solving a performance driven optimization problem, the generalized likelihood ratio test (GLRT) is better looked at as a heuristic principle than as a test strategy having assured optimality properties. However, as will be discussed below, the GLRT is a straightforward procedure and it does have asymptotic (large n) optimality properties that are major attractions.

We consider the general composite hypotheses

$$\begin{aligned} H_0 &: \theta \in \Theta_0 \\ H_1 &: \theta \in \Theta_1 \end{aligned}$$

The GLRT can be defined as an "estimate-and-plug" procedure to test H_0 vs. H_1:

Step 1: Find good estimates $\hat{\theta}_0$ and $\hat{\theta}_1$ of θ under H_0 and H_1

Step 2: substitute these estimates into the LR statistic

Using this procedure we obtain the GLRT

$$\Lambda = \frac{f(x;\hat{\theta}_1)}{f(x;\hat{\theta}_0)} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

where η is selected to give FA level α

Any consistent estimators will ensure that the GLRT has favorable asymptotic properties. However, the most common case of the GLRT is when the estimators $\hat{\theta}_1$ and $\hat{\theta}_0$ are MLEs:

$$\Lambda_{\text{GLR}} = \frac{\max_{\theta\in\Theta_1} f(x;\theta)}{\max_{\theta\in\Theta_0} f(x;\theta)} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

Note: these MLE's are constrained to $\theta \in \Theta_0$ and $\theta \in \Theta_1$, respectively.

For a simple hypothesis H_0 the GLRT reduces to

$$\begin{aligned} \Lambda_{\text{GLR}} &= \frac{\sup_{\theta\in\Theta_1} f(x;\theta)}{f(x;\theta_0)} \\ &= \sup_{\theta\in\Theta_1} \Lambda(\theta) = \Lambda(\hat{\theta}_1). \end{aligned}$$

where

$$\Lambda(\theta) = \frac{f(x;\theta)}{f(x;\theta_0)}$$

8.10.1 PROPERTIES OF GLRT

The following properties are stated simply and without proof but can be expected to hold for smooth likelihood functions and a simple null hypothesis. For proofs of these properties the reader is referred to [40] or [7].

1. If an UMP test exists then the GLRT will be identical to it.

2. Let observations $\underline{X} = [X_1, \ldots, X_n]^T$ be i.i.d. Then, since the MLE $\hat{\theta}$ is a consistent estimator, as $n \to \infty$ the GLRT is asymptotically UMP.

3. The GLR test statistic for testing a double sided alternative hypothesis has a Chi-square limiting distribution under H_0 as $n \to \infty$. Specifically, assume that the unknown parameters are partitioned as

$$\underline{\theta} = [\varphi_1, \ldots, \varphi_p, \quad \underbrace{\xi_1, \ldots, \xi_q}_{\text{nuisance parameters}} \quad]^T$$

and consider the GLRT for the simple null and double-sided alternative hypotheses

$$\begin{aligned} H_0 &: \underline{\varphi} = \underline{\varphi}_0, \quad \underline{\xi} = \underline{\xi}_o \\ H_1 &: \underline{\varphi} \neq \underline{\varphi}_0, \quad \underline{\xi} = \underline{\xi}_o \end{aligned}$$

where $\underline{\xi}_o$ is unknown. Then for large n

$$2 \ln \Lambda_{\text{GLR}}(X) \ \sim \ \chi_p, \quad \text{under } H_0. \tag{113}$$

Note that p is the number of parameters that are unknown under H_1 but are fixed under H_0.

8.11 BACKGROUND REFERENCES

Any of the references cited in the last chapter will have some discussion of the problem of testing composite hypotheses. Lehmann [39] has comprehensive coverage of minimax, similarity (CFAR), unbiased tests, and other methods. Invariance principles have been applied to many problems in signal processing and communications [37],[10], [9], [61]. The book by Kariya and Sinha [33] is a comprehensive, advanced level, reference on invariance principles, robustness and GLR's (therein referred to as likelihood principles) relevant to these studies. Application of invariance principles can be viewed as one way to figure out a transformation of the data that makes the resultant transformed measurements have more tractable density functions under H_0 or H_1. Viewed in this way these principles can be interpreted as a special application of the *transformation method*, discussed in the context of robust exploratory data analysis in the book by Hoaglin, Mosteller and Tukey [27]. Another approach, that we did not discuss here, is the application of non-parametric techniques, also called "distribution-free inference," to handle unknown parameters in testing of composite hypotheses. The book by Hollander and Wolfe [28] covers this topic from a general statistical point of view and the edited book by Kassam and Thomas [35] covers nonparametric detection theory for applications in signal processing, communications and control.

8.12 EXERCISES

8.1 The observations $\{x_i\}_{i=1}^n$ are i.i.d. exponential $x_i \sim f_\theta(x) = \beta e^{-\beta x}$, where $x, \beta \geq 0$. Consider testing the following single sided hypotheses

$$\begin{aligned} H_0 &: \beta = \beta_0 \\ H_1 &: \beta > \beta_0 \end{aligned}$$

(a) First find the MP test of level α for the simple alternative $H_1 : \beta = \beta_1$ where $\beta_1 > \beta_0$. Express the threshold in terms of the Gamma distribution (distribution of n i.i.d. exponential r.v.s). Next establish that your test is UMP for the single sided composite H_1 above.

(b) Specialize the results of (a) to the case of a single observation $n = 1$ and derive the ROC curve. Plot your curve for $\beta_1/\beta_0 = 1, 5, 10$.

(c) Derive the locally most powerful test (LMPT) for the single sided hypotheses (maximize slope of power subject to FA constraint) and verify that it is identical to the UMP test.

(d) Now consider testing the double sided hypotheses

$$H_0 : \beta = \beta_0$$
$$H_1 : \beta \neq \beta_0$$

Derive the LMPT (maximize curvature of power subject to FA constraint and zero slope condition). Derive the ROC for $n = 1$ and compare to the ROC of part (b) over the region $\beta > \beta_0$.

(e) Derive the GLRT for the double sided hypotheses of part (d). Compare to your answer obtained in part (d).

8.2 Let Z be a single observation having density function

$$p_\theta(z) = (2\theta z + 1 - \theta), \qquad 0 \leq z \leq 1$$

where $-1 \leq \theta \leq 1$.

(a) Is there a uniformly most powerful test between the composite hypotheses

$$H_0 : \theta = 0$$
$$H_1 : \theta \neq 0$$

and, if so, what is it?

(b) Find the generalized likelihood ratio test for these hypotheses.

(c) Now assume that under H_1 the parameter θ has prior density $p(\theta) = |\theta| I_{[-1,1]}(\theta)$ so that under H_1 the density of Z is $f_1(z) = \int f(z|\theta) p_\theta(\theta) d\theta$, where $f(z|\theta) = p_\theta(z)$. Find the MP test between hypotheses H_0 and this new H_1. What if the prior density were the assymetric $p(\theta) = \frac{1}{2}(\theta + 1) I_{[-1,1]}(\theta)$?

8.3 A random variable X has density

$$f(x; \theta) = \frac{1 + \theta x}{2}, \quad -1 \leq x \leq 1$$

where $\theta \in [-1, 1]$.

(a) Find the MP test of level α for testing the simple hypotheses

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_1 &: \theta = \theta_1 \end{aligned}$$

based on a single sample x, where $\theta_0 \in [-1, 0]$ and $\theta_1 \in (0, 1]$ are known. Derive and plot the ROC when $\theta_0 = 0$.

(b) Is there a UMP test of level α, and if so what is it, for the following hypotheses?

$$\begin{aligned} H_0 &: \theta = 0 \\ H_1 &: \theta > 0 \end{aligned}$$

(c) Now consider testing the doubly composite hypotheses

$$\begin{aligned} H_0 &: \theta \leq 0 \\ H_1 &: \theta > 0 \end{aligned}$$

Find the GLRT for the above hypotheses. Derive the threshold of the GLRT that ensures the level α condition $\max_{\theta\in[-1,0]} P_{FA}(\theta) \leq \alpha$.

8.4 Available is an i.i.d. sample of a Poisson r.v. with distribution $p_\theta(k) = P_\theta(x_i = k) = \frac{\theta^k}{k!}e^{-\theta}$, $k = 0, 1, 2, \ldots$.

(a) Find the GLRT for testing the hypotheses

$$\begin{aligned} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{aligned}$$

Do not attempt to set the exact threshold for level α.
In the following parts of this exercise you will show how to set the GLRT threshold under the large n Chi-square approximation to the GLRT test statistic $\Lambda = max_{\theta\neq\theta_o} p_\theta(\underline{x})/p_{\theta_o}(\underline{x})$.

(b) Directly show that under H_0 the statistic $2\log\Lambda$ is asymptotically Chi-square with 1 d.f. by expanding $\Lambda = \Lambda(\overline{x_i})$ about the sample mean $\overline{x_i} = \theta_o$, neglecting all terms of order $(\overline{x_i} - \theta_o)^3$ and higher, and recalling that $(\mathcal{N}(0,1))^2$ is Chi-square with 1 d.f.

(c) Using the result of part (b) set the threshold of your GLRT in part (a).

(d) Using the asymptotic results of part (b) find the GLRT between

$$\begin{aligned} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{aligned}$$

with threshold.

8.5 Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with the marginal density $X_i \sim f(x) = \epsilon g(x) + (1-\epsilon)h(x)$, where $\epsilon \in [0,1]$ is a non-random constant and $g(x)$ and $h(x)$ are known density functions. It is desired to test the composite hypotheses

$$H_0 : \epsilon = 1/2 \tag{114}$$
$$H_1 : \epsilon > 1/2 \tag{115}$$

(a) Find the most powerful (MP) test between H_0 and the simple hypothesis $H_1 : \epsilon = \epsilon_1$, where $\epsilon_1 > 1/2$ (you needn't solve for the threshold). Is your MP test a UMP test of the composite hypotheses (115)?

(b) Find the locally most powerful (LMP) test for (115). Show how you can use the CLT to set the threshold for large n.

(c) Find the generalized LRT (GLRT) test for (115) in the case of $n = 1$. Compare to your answer in part (b).

8.6 Let $\{X_i\}_{i=1}^n$ be i.i.d. following an exponential distribution

$$f(x;\theta) = \theta e^{-\theta x}, \quad x \geq 0$$

with $\theta > 0$. You are to design a test of the hypotheses

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_1 &: \theta \neq \theta_0 \end{aligned}$$

Here we explore various testing strategies.

(a) Show that the GLRT reduces to a test on the sum of the X_i's and derive the threshold to attain FA level α (Hint: the sum of n standard (mean = 1) exponential r.v.s is standard Gamma with parameter n).

(b) Now assume that H_0 and H_1 have equal prior probabilities $p = 1/2$ and that, conditioned on H_1, θ itself follows an exponential distribution of the form $f(\theta) = \beta e^{-\beta\theta}$, $\theta \geq 0$, where $\beta > 0$ is known. Find the form of the Bayes LRT (with threshold) which attains minimum probability of decision error. What happens as $\beta \to \infty$?.

8.7 As in Exercise 4.24, let n i.i.d. realizations be available from the geometric mixture $f_{\mathcal{G}}$ specified by (57) and (58). Assume that ϕ_1, ϕ_2 are known and $\phi_1 \neq \phi_2$.

(a) Consider the hypothesis testing problem on $f_{\mathcal{G}}$

$$\begin{aligned} H_0 &: \epsilon = 0 \\ H_1 &: \epsilon > 0. \end{aligned}$$

Does a level α UMP test for these hypotheses exist? If so what is it? If not derive a GLRT test. You must specify a threshold of level α (you can assume large n).

(b) Consider the hypothesis testing problem on $f_{\mathcal{G}}$

$$\begin{aligned} H_0 &: \epsilon = 1/2 & (116) \\ H_1 &: \epsilon \neq 1/2. & (117) \end{aligned}$$

Does a level α UMP test for these hypotheses exist? If so what is it? If not derive a GLRT test. You must specify a threshold of level α (you can assume large n).

(c) Under the identical assumptions as in part (h) find a locally most powerful unbiased test of (117) based on n i.i.d. observations from $f_{\mathcal{G}}$ and compare to the GLRT.

8.8 Let X be a random variable with density $f(x;\theta) = (\theta+1)x^\theta$, $x \in [0,1]$ and $\theta > -1$. Consider testing the hypotheses

$$\begin{aligned} H_0 &: \theta = 0 \\ H_1 &: \theta = \theta_1 \end{aligned} \qquad (118)$$

(a) Find the most powerful (MP) test of level α for testing these hypotheses and derive expressions for the power function and ROC curve. Does the decision region of the MP test of level α depend on the value of θ_1?. Does there exist a UMP test of level α for testing H_0 vs. $H_1 : \theta > 0$? How about for testing H_0 against $H_1 : \theta \neq 0$?

(b) Assuming priors on the hypotheses $p = P(H_0)$, $1-p = P(H_1)$ find the optimal Bayes test of (118) under the assumption that $c_{00} = c_{11} = 0$ and $c_{01} = c_{10} = 1$ (minimal probability of error test). Find and plot the minimum risk (probability of error) $\overline{c}^*(p)$ as a function of p for $\theta_1 = 1$. Using these results find the mini-max Bayes detector and its threshold for this value of θ.

(c) Find the locally most powerful test for testing H_0 vs. $H_1 : \theta > 0$ and derive an expression for the ROC curve.

(d) Find the GLRT for testing H_0 against $H_1 : \theta \neq 0$ and derive expressions for P_F and P_D in terms of the threshold and plot the ROC curve.

8.9 In this exercise you will explore the problem of detecting an anomaly in an image solely on the basis of filtered measurements, e.g. a blurred version of the image. This type of problem is related to "non-destructive testing" and arises in many situations that we encounter in our daily lives, e.g., when we pass our suitcases through a security scanner at the airport. When there is no noise and no anomaly the scanner outputs an image that lies in a known subspace, the span of the columns of a known $n \times p$ matrix $\mathbf{H}$, denoted colspan$\{\mathbf{H}\}$. You might just think of the columns of $\mathbf{H}$ as blurry images of all the possible "benign" objects that one could pack into a suitcase. An anomaly occurs when the image has components lying outside of this subsapce of benign objects. Of course, there is also additive noise that complicates our ability to detect such anomalies.

Now we can state the anomaly detection problem as testing the hypotheses

$$\begin{aligned} H_0 &: \underline{X} = \mathbf{H}\underline{\theta} + \underline{W} \\ H_1 &: \underline{X} = \underline{\psi} + \mathbf{H}\underline{\theta} + \underline{W}, \end{aligned} \tag{119}$$

where we have defined the observed image as a vector $\underline{X} = [X_1, \ldots, X_n]^T$, the parameter vector $\underline{\theta} = [\theta_1, \ldots, \theta_p]^T$ describes the specific linear combination of benign objects present in the suitcase, $\underline{\psi} = [\psi_1, \ldots, \psi_n]^T$ describes the anomalous component of the image, and $\underline{W}$ is a Gaussian noise vector with zero mean and covariance matrix $\text{cov}(\underline{W}) = \sigma^2\mathbf{I}$. We will assume throughout this exercise that we know the matrix $\mathbf{H}$ and σ^2. We also assume that $\mathbf{H}$ is full rank: $\text{rank}(\mathbf{H}) = p \leq n$.

(a) Assume that $\underline{\theta}$ is known. For known $\underline{\psi}$ what is the most powerful (MP) test of level α for testing H_0 vs. H_1? Is the test you derived in part (a) UMP for testing H_0 vs. $H_1 : \underline{X} = c\underline{\psi} + \mathbf{H}\underline{\theta} + \underline{W}$, where $c > 0$ is an unknown constant? Is the test UMP for totally unknown $\underline{\psi}$?

(b) Find an expression for and plot the ROC curve (hand drawn is fine) for the test derived in (a). What function of $\underline{\psi}$, $\underline{\theta}$, $\mathbf{H}$, and σ determines the shape of the ROC, i.e., detectibility index?

(c) Now assume that $\underline{\theta}$ is unknown but that $\underline{\psi}$ is known. Find the GLRT of level α for testing (119) and find its ROC curve. What function of $\underline{\psi}$, $\underline{\theta}$, $\mathbf{H}$, and σ determines the shape of the ROC, i.e., detectibility index? What happens to the detectibility when $p = n$?

(d) Now assume that $\underline{\theta}$ and $\underline{\psi}$ are both unknown. Find the GLRT for testing (119).

(e) Assume that $\underline{\psi}$ is known but $\underline{\theta}$ is unknown. Also assume that the anomaly vector satisfies the constraint $\underline{\psi}^T\underline{\psi} \leq \epsilon$, $\epsilon > 0$. Using the results you derived in (c) find the least detectable (giving lowest power) and the most detectable (giving highest power) anomaly vectors.

8.10 Assume X is a Cauchy distributed random variable with density

$$f(x;\theta) = \frac{1}{\pi}\frac{1}{1 + (x-\theta)^2}.$$

You are to test the hypotheses $H_0 : \theta = 0$ vs. $H_1 : \theta > 0$ based on a single realization of X.

(a) Derive the MP test of level α for testing $H_0 : \theta = 0$ vs. $H_1 : \theta = \theta_1$ for a fixed value $\theta_1 > 0$.

(b) Find the decision region $\mathcal{X}_1$ specifying outcomes X that will result in deciding H_1.

(c) Show that this decision region depends on θ_1 and therefore establish that no UMP exists.

8.11 In many applications it is of interest to detect deviations of a random sample from a nominal probability model. Such deviations are called anomalies and the general problem can be formulated as a hypothesis testing problem. One popular test uses "minimum volume sets," and is explored in this exercise. In the following X is a single realization of a measurement vector in $\mathbb{R}^d$ to be tested for anomaly.

Assume that a nominal density $f_0(x)$ for X is known, e.g., learned from a lot of (non-anomalous) training samples. We assume that $\int_{[0,1]^d} f_0(x)dx = 1$, that is, f_0 is supported on the d-dimensional unit cube $[0,1]^d$. We hypothesize that an anomalous observation has a density f_1 that corresponds to a broadening of f_0. Arguably, the simplest model is that f_1 is the superposition of f_0 and a uniform density over $[0,1]^d$:

$$f_1(x) = (1-\epsilon)f_0(x) + \epsilon U(x)$$

where

$$U(x) = \begin{cases} 1, & x \in [0,1]^d \\ 0, & o.w. \end{cases}$$

and ϵ is a mixture parameter $0 < \epsilon \leq 1$. With this model we can define the pdf of X as $f(x) = (1-\epsilon)f_0(x) + \epsilon U(x)$ and say that a nominal measurement corresponds to $\epsilon = 0$ while an anomaly corresponds to $\epsilon > 0$.

(a) First we consider the case that the anomalous density is known, i.e., $\epsilon = \epsilon_1$, $\epsilon > 0$. Show that the most powerful (MP) test of level α of the "simple anomaly hypothesis"

$$\begin{aligned} H_0 &: \epsilon = 0 \\ H_1 &: \epsilon = \epsilon_1 \end{aligned}$$

is of the form "decide H_1" if

$$f_0(X) \leq \eta$$

where η is a suitably selected threshold.

(b) Show that the MP test of part (a) is a minimum-volume-set in the sense that the H_0-decision region $\Omega^* = \{x : f_0(x) \geq \eta\}$ is a set having minimum volume over all sets Ω satisfying $\int_\Omega f_0(x)dx = 1-\alpha$. (Hint: express the volume of Ω as $|\Omega| = \int \phi(x)dx$, where $\phi(x)$ is the indicator function of the set Ω).

(c) Derive the power function of the minimum-volume-set test and show that it is proportional to the volume of Ω.

(d) Next we consider the case that the anomalous density is unknown. Is the minimum-volume-set test of part (b) uniformly most powerful (UMP) for the "composite anomaly hypothesis"

$$\begin{aligned} H_0 &: \epsilon = 0 \\ H_1 &: \epsilon > 0 \end{aligned}$$

If it is not UMP derive the locally most powerful test (LMP) of level α.

(e) Now specialize to the case of a scalar observation X, i.e., $d = 1$, where $f_0(x)$ is the triangular density

$$f_0(x) = \begin{cases} 4x, & 0 \leq x \leq 1/2 \\ 4(1-x), & 1/2 < x \leq 1 \\ 0, & o.w. \end{cases}$$

Find mathematical expressions for the set Ω^* in terms of the false alarm level α, derive the power function (as a function of ϵ), and the ROC curve. Plot the power function for several representative values of α and plot the ROC curve for several representative values of ϵ.

End of chapter

9 COMPOSITE HYPOTHESES IN THE UNIVARIATE GAUSSIAN MODEL

In this chapter we illustrate the generalized likelihood ratio testing strategy discussed in Chapter 8 to hypotheses on the mean and variance of the univariate Gaussian distribution based on i.i.d measurements. We will deal with the following scenarios:

* Tests on mean of a single population: σ^2 known
* Tests on mean of a single population: σ^2 unknown
* Tests on variance of a single population: μ known
* Tests on variance of a single population: μ unknown
* Tests on equality of means in two populations
* Tests on equality of variances in two populations
* Tests on correlation between two populations

Recall the form of the density of an i.i.d. Gaussian vector $\underline{X} = [X_1, \ldots, X_n]^T$ with mean μ and variance σ^2.

$$f(\underline{x};\ \mu, \sigma) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

9.1 TESTS ON THE MEAN: σ^2 KNOWN

Case I: $H_0 : \mu = \mu_o$, $H_1 : \mu > \mu_o$

Case II: $H_0 : \mu \leq \mu_o$, $H_1 : \mu > \mu_o$

Case III: $H_0 : \mu = \mu_o$, $H_1 : \mu \neq \mu_o$

We have already established that UMP test exists for Case I. You can show that same test is UMP for case II by checking the monotone likelihood condition [39] discussed in Sec. 8.2.

9.1.1 CASE III: $H_0 : \mu = \mu_o$, $H_1 : \mu \neq \mu_o$

$\underline{X} = [X_1, \ldots, X_n]^T$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$

Here $\theta = \mu$, $\Theta = \mathbb{R}$ and we want to test the double sided hypothesis that the mean equals a known parameter μ_0

$$\begin{aligned} H_0 : \mu &= \mu_o \\ H_1 : \mu &\neq \mu_o. \end{aligned}$$

The GLRT is of the form:

$$\Lambda_{\text{GLR}} = \max_{\mu\neq\mu_o}\Lambda(\mu) = \frac{\max_{\mu\neq\mu_o}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right)}{\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu_o)^2\right)}$$

We must consider two cases

1. $\hat{\mu}_{ml} \neq \mu_o$

2. $\hat{\mu}_{ml} = \mu_o$

Case 1. $\hat{\mu}_{ml} \neq \mu_o$:

In this case it is obvious that

$$\max_{\mu \neq \mu_o} \Lambda(\mu) = \Lambda(\hat{\mu}_{ml})$$

Case 2. $\hat{\mu}_{ml} = \mu_o$:

Since $\Lambda(\mu)$ is a continuous function with maximum at $\mu = \hat{\mu}_{ml}$ we have again

$$\max_{\mu \neq \mu_o} \Lambda(\mu) = \lim_{\epsilon \to 0} \Lambda(\hat{\mu}_{ml} + \epsilon) = \Lambda(\hat{\mu}_{ml})$$

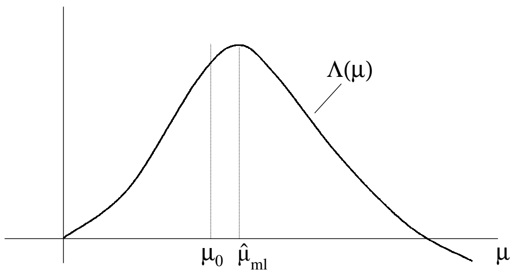


Figure 124: *For a continuous LR density $f(x;\mu)$ the maximum of the likelihood ratio test statistic $\lambda(\mu)$ occurs at the MLE $\mu = \hat{\mu}_{ml}$*

Thus, since we know $\hat{\mu}_{ml}$ is the sample mean under the Gaussian model

$$\Lambda_{\text{GLR}} = \frac{\exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^n (X_j - \overline{X})^2\right)}{\exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^n (X_j - \mu_o)^2\right)}$$

Next use the fact that $\sum_{i=1}^n (X_i - \overline{X}) = 0$ (recall that $\overline{X}$ is the LLS estimator over all estimator functions that are independent of the data) to obtain

$$\sum_{j=1}^n (X_j - \mu_o)^2 = \sum_{j=1}^n (X_j - \overline{X})^2 + n(\overline{x} - \mu_o)^2.$$

Hence,

$$\Lambda_{\rm GLR} = \exp\left(\frac{n}{2\sigma^2}(\overline{X}-\mu_o)^2\right)$$

and the GLRT is simply

$$\frac{\sqrt{n}\ \left|\overline{X}-\mu_o\right|}{\sigma} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma = \mathcal{N}^{-1}(1-\alpha/2), \tag{120}$$

which is identical to the LMPU (UMPU) test derived in Sec. 8.6!

Note: as predicted by our results on the asymptotic distribution of GLRTs of double sided hypotheses (Sec. 8.10.1)

$$\begin{aligned} 2\ln\Lambda_{\rm GLR} &= 2\ln\left\{\exp\left(\frac{n}{2\sigma^2}(\overline{X}-\mu_o)^2\right)\right\} \\ &= \left(\underbrace{\frac{\overline{X}-\mu_o}{\sigma/\sqrt{n}}}_{\mathcal{N}(0,1)}\right)^2 \end{aligned}$$

which is distributed as a central Chi-square with 1 d.f.

A general lesson learned for GLRT's:

⇒ for testing double sided hypotheses of form

$$\begin{aligned} H_0 : \underline{\theta} &= \underline{\theta}_o \\ H_1 : \underline{\theta} &\neq \underline{\theta}_o \end{aligned}$$

if LR $\Lambda(\underline{\theta})$ is a continous function of $\underline{\theta}$ then

$$\max_{\underline{\theta}\neq\underline{\theta}_o} \Lambda(\underline{\theta}) = \max_{\underline{\theta}} \Lambda(\underline{\theta}) = \Lambda(\underline{\hat{\theta}}_{ml})$$

9.2 TESTS ON THE MEAN: σ^2 UNKNOWN

Case I: $H_0 : \mu = \mu_o,\ \sigma^2 > 0,\ H_1 : \mu > \mu_o,\ \sigma^2 > 0$

Case II: $H_0 : \mu \leq \mu_o,\ \sigma^2 > 0,\ H_1 : \mu > \mu_o,\ \sigma^2 > 0$

Case III: $H_0 : \mu = \mu_o,\ \sigma^2 > 0,\ H_1 : \mu \neq \mu_o,\ \sigma^2 > 0$

9.2.1 CASE I: $H_0 : \mu = \mu_o,\ \sigma^2 > 0$, $H_1 : \mu > \mu_o,\ \sigma^2 > 0$

From properties of the MLE for Gaussian mean and variance parameters we can easily show that for a realization $\underline{x}$ of $\underline{X}$

$$\begin{aligned} \Lambda_{\rm GLR} &= \frac{\max_{\mu>\mu_o,\sigma^2>0} f(\underline{x};\mu,\sigma^2)}{\max_{\sigma^2>0} f(\underline{x};\mu_o,\sigma^2)} \\ &= \begin{cases} \frac{f(\underline{x};\,\overline{x},\,\overline{(x_i-\overline{x})^2})}{f(\underline{x};\,\mu_o,\,\overline{(x_i-\mu_o)^2})}, & \overline{x} > \mu_o \\ 1, & \overline{x} \leq \mu_o \end{cases} \end{aligned}$$

where $f(x;\mu,\sigma^2)$ is the $\mathcal{N}(\mu,\sigma^2)$ density and (as usual)

$$\begin{aligned}\overline{x} &= n^{-1}\sum_{i=1}^n x_i \\ \overline{(x_i-t)^2} &= n^{-1}\sum_{i=1}^n (x_i-t)^2 = \hat{\sigma}_t^2.\end{aligned}$$

Here t is an arbitrary constant.

Next observe that

$$\begin{aligned}\frac{f(\underline{x};\ \overline{x},\ \overline{(x_i-\overline{x})^2})}{f(\underline{x};\ \mu_o,\ \overline{(x_i-\mu_o)^2})} &= \left(\frac{\overline{(x_i-\mu_o)^2}}{\overline{(x_i-\overline{x})^2}}\right)^{n/2} \\ &= \left(1+\frac{(\overline{x}-\mu_o)^2}{\overline{(x_i-\overline{x})^2}}\right)^{n/2} \\ &= \left(1+T^2(\underline{x})\right)^{n/2},\end{aligned}$$

where

$$T(\underline{x}) = \frac{\overline{x}-\mu_o}{\sqrt{\overline{(x_i-\overline{x})^2}}}$$

Since $T^2(\underline{x})$ is monotone in $T(\underline{x})$ for $\overline{x}>\mu_o$ the GLRT based on $\underline{X}$ is

$$T(\underline{X}) = \frac{\overline{X}-\mu_o}{\sqrt{\overline{(X_i-\overline{X})^2}}} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

which is equivalent to the one sided t-test:

$$\frac{(\overline{X}-\mu_o)}{\mathsf{s}/\sqrt{n}} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma'$$

where recall that s^2 is the *unbiased* variance estimate for unknown mean

$$\mathsf{s}^2 = (n-1)^{-1}\sum_{i=1}^n (X_i-\overline{X})^2$$

PERFORMANCE:

Under H_0 we have

$$\begin{aligned}T(\underline{X}) &= \frac{\overbrace{\sqrt{n}\,(\overline{X}_i-\mu_o)}^{\mathcal{N}(0,1)\cdot\sigma}}{\sqrt{\underbrace{\sum_{i=1}^n (X_i-\overline{X}_i)^2}_{\chi_{n-1}\cdot\sigma^2}/(n-1)}} \\ &= \frac{\mathcal{N}(0,1)}{\sqrt{\chi_{n-1}/(n-1)}} = \mathcal{T}_{n-1}\end{aligned}$$

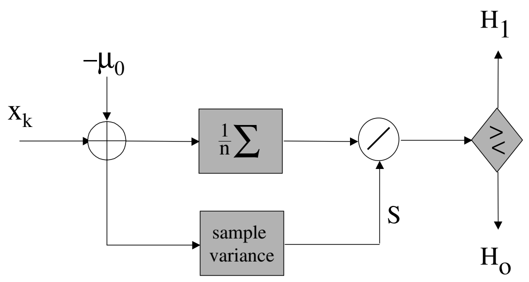


Figure 125: *The one sided t-test for detection of mean exceeding μ_o when variance is unknown*

where $\mathcal{T}_{n-1}$ is Student-t r.v. with $n-1$ d.f. Thus

$$\alpha = P_0(T(\underline{X}) > \gamma) = 1 - \mathcal{T}_{n-1}(\gamma),$$

so that

$$\gamma = \mathcal{T}_{n-1}^{-1}(1-\alpha).$$

Under H_1 we have:
* $\overline{X_i} - \mu_o$ has mean $\mu - \mu_o$ which is no longer zero.
$\Rightarrow T(\underline{X})$ follows the non-central Student-t distribution
$\mathcal{T}_{n,d}$ with $n-1$ d.f. and non-centrality parameter

$$d = \frac{\sqrt{n}(\mu_1 - \mu_0)}{\sigma}.$$

Hence, the power of the one sided t-test is

$$\beta = 1 - \mathcal{T}_{n-1,d}(\mathcal{T}_{n-1}^{-1}(1-\alpha)),$$

which is plotted in Fig. 126.

Note: for large n, $\mathcal{T}_{n-1,d} \rightarrow \mathcal{N}_1(d,1)$ and therefore the power function approaches the power function of the GLRT for single sided hypothesis on the mean with σ^2 known.

9.2.2 CASE II: $H_0 : \mu \leq \mu_o,\ \sigma^2 > 0$, $H_1 : \mu > \mu_o,\ \sigma^2 > 0$

We can show that the GLRT has identical one-sided t-test form as in Case I. (Exercise)

$$\frac{\sqrt{n}\ (\overline{x} - \mu_o)}{s} \overset{H_1}{\underset{H_0}{\gtrless}} \mathcal{T}_{n-1}^{-1}(1-\alpha)$$

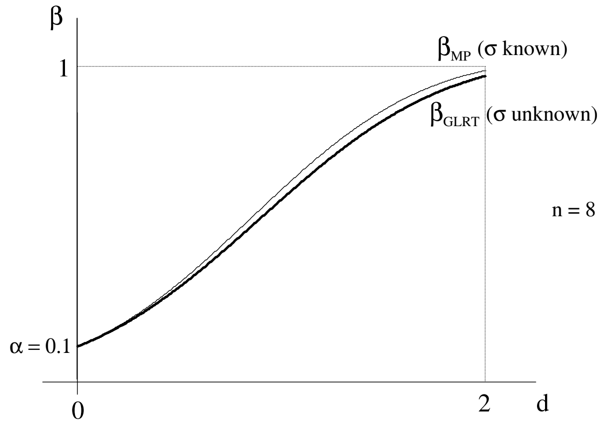


Figure 126: *Power curve for one sided t-test*

9.2.3 CASE III: $H_0 : \mu = \mu_o,\ \sigma^2 > 0$, $H_1 : \mu \neq \mu_o,\ \sigma^2 > 0$

This case is very similar to the double sided GLRT on the mean for known σ^2. We obtain GLRT as double sided t-test

$$\frac{\sqrt{n}|\overline{x} - \mu_o|}{s} \overset{H_1}{\underset{H_0}{\gtrless}} \mathcal{T}_{n-1}^{-1}(1 - \alpha/2)$$

with power curve

$$\beta = 1 - \mathcal{T}_{n-1,d}(\mathcal{T}_{n-1}^{-1}(1 - \alpha/2)) + \mathcal{T}_{n-1,-d}(\mathcal{T}_{n-1}^{-1}(1 - \alpha/2)).$$

For large n this converges to the power curve derived in the case of known variance.

9.3 TESTS ON VARIANCE: KNOWN MEAN

CASE I: $H_0 : \sigma^2 = \sigma_o^2$, $H_1 : \sigma^2 > \sigma_o^2$
CASE II: $H_0 : \sigma^2 \leq \sigma_o^2$, $H_1 : \sigma^2 > \sigma_o^2$
CASE III: $H_0 : \sigma^2 = \sigma_o^2$, $H_1 : \sigma^2 \neq \sigma_o^2$

9.3.1 CASE I: $H_0 : \sigma^2 = \sigma_o^2$, $H_1 : \sigma^2 > \sigma_o^2$

Similarly to the derivation of the GLRT for the case of one sided tests of the mean (120), the continuity of the Gaussian p.d.f. $f(\underline{x}; \mu, \sigma^2)$ as a function of σ^2 gives:

$$\begin{aligned}
\Lambda_{\mathrm{GLR}} &= \frac{\max_{\sigma^2 > \sigma_o^2} f(\underline{x}; \mu, \sigma^2)}{f(\underline{x}; \mu, \sigma_o^2)} \\
&= \begin{cases} \frac{f(\underline{x}; \mu, \hat{\sigma}_\mu^2)}{f(\underline{x}; \mu, \sigma_o^2)}, & \hat{\sigma}_\mu^2 > \sigma_o^2 \\ 1, & \hat{\sigma}_\mu^2 \leq \sigma_o^2 \end{cases}
\end{aligned}$$

where $\hat{\sigma}_\mu^2$ is the *unbiased* estimate of the variance for known mean

$$\hat{\sigma}_\mu^2 = n^{-1} \sum_{i=1}^n (x_i - \mu)^2$$

After some simple manipulation the GLRT takes the form

$$\Lambda_{\mathrm{GLR}} = \left(\underbrace{\frac{1}{\max\{\hat{\sigma}_\mu^2/\sigma_o^2,\ 1\}}}_{u} e^{\max\{\hat{\sigma}_\mu^2/\sigma_o^2,\ 1\} - 1} \right)^{n/2} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \eta$$

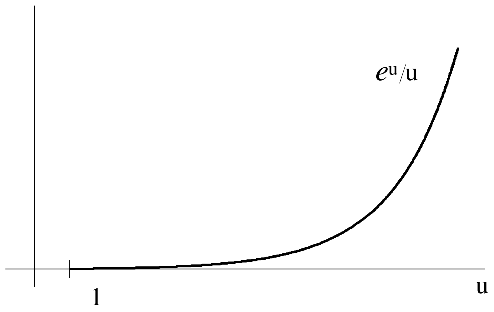


Figure 127: *The function e^u/u is monotone increasing over $u \geq 1$.*

As the function e^u/u is monotone increasing over $u \geq 1$, the GLRT reduces to

$$\max\{\hat{\sigma}_\mu^2/\sigma_o^2, 1\} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \gamma$$

Now if $\gamma \leq 1$ then false alarm $\alpha = 1$.
Hence we can select $\gamma > 1$ and GLRT reduces to the single sided Chi-square test

$$T(\underline{x}) = \frac{n\hat{\sigma}_\mu^2}{\sigma_o^2} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \chi_n^{-1}(1 - \alpha)$$

which we know from previous work is actually an UMP test.

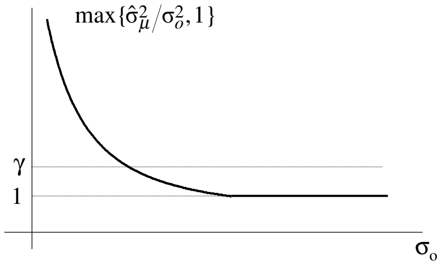


Figure 128: *The GLRT always chooses H_1 for $\gamma < 1$.*

An alternative form of the UMP test is a single sided energy detector

$$\sum_{i=1}^{n}(x_i - \mu)^2 \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \gamma$$

9.3.2 CASE II: $H_0 : \sigma^2 \leq \sigma_o^2,\ H_1 : \sigma^2 > \sigma_o^2$

Now we have

$$\Lambda_{\text{GLR}} = \frac{\max_{\sigma^2 > \sigma_o^2} f(\underline{x};\ \mu,\ \sigma^2)}{\max_{\sigma^2 \leq \sigma_o^2} f(\underline{x};\ \mu,\ \sigma^2)}$$

$$= \begin{cases} \frac{f(\underline{x};\,\mu,\,\hat{\sigma}_\mu^2)}{f(\underline{x};\,\mu,\,\sigma_o^2)}, & \hat{\sigma}_\mu^2 > \sigma_o^2 \\ \\ \frac{f(\underline{x};\,\mu,\,\sigma_0^2)}{f(\underline{x};\,\mu,\,\hat{\sigma}_\mu^2)}, & \hat{\sigma}_\mu^2 \leq \sigma_o^2 \end{cases}$$

or

$$\Lambda_{\text{GLR}} = \begin{cases} \left(\frac{1}{\hat{\sigma}_\mu^2/\sigma_o^2}\, e^{\hat{\sigma}_\mu^2/\sigma_o^2 - 1}\right)^{n/2}, & \hat{\sigma}_\mu^2 > \sigma_o^2 \\ \left(\hat{\sigma}_\mu^2/\sigma_o^2\, e^{1-\hat{\sigma}_\mu^2/\sigma_o^2}\right)^{n/2}, & \hat{\sigma}_\mu^2 \leq \sigma_o^2 \end{cases}$$

As e^u/u is monotone increasing over $u > 1$ and ue^{-u} is monotone increasing over $0 \leq u \leq 1$, the GLRT reduces to the same form as derived for Case I:

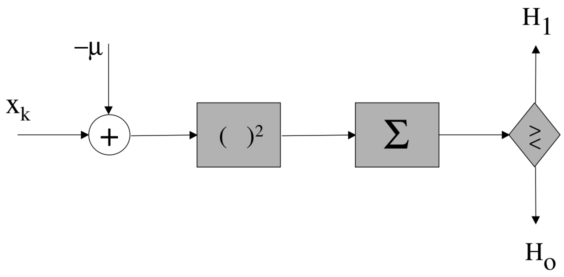


Figure 129: *GLRT for one sided test of positive shift in variance is an energy detector.*

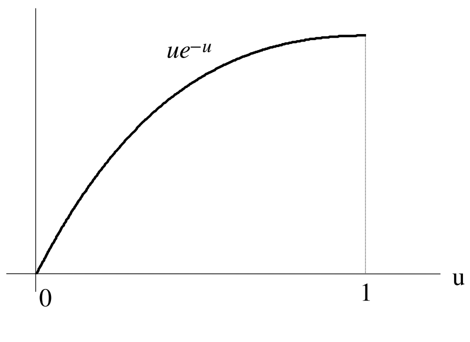


Figure 130: *The function ue^{-u} is monotone increasing over $0 \leq u \leq 1$.*

$$\frac{n\,\hat{\sigma}_\mu^2}{\sigma_o^2} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

and the rest of the analysis is identical to before.

9.3.3 CASE III: $H_0 : \sigma^2 = \sigma_o^2,\ H_1 : \sigma^2 \neq \sigma_o^2$

Now, we have

$$\begin{aligned} \Lambda_{\mathrm{GLR}} &= \frac{\max_{\sigma^2 \neq \sigma_o^2} f(\underline{x}; \mu, \sigma^2)}{f(\underline{x}; \mu, \sigma_o^2)} \\ &= \frac{f(\underline{x}; \mu,\ \hat{\sigma}_\mu^2)}{f(\underline{x};\ \mu,\ \sigma_o^2)} \\ &= \left(\frac{1}{\hat{\sigma}_\mu^2/\sigma_o^2}\ e^{\hat{\sigma}_\mu^2/\sigma_o^2 - 1}\right)^{n/2} \end{aligned}$$

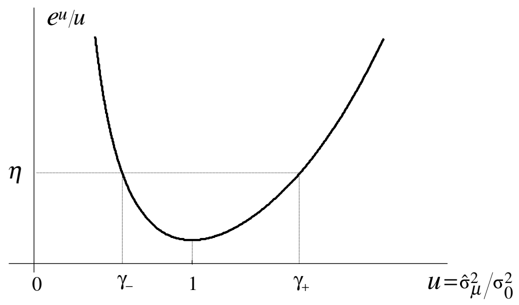


Figure 131: *As e^u/u is convex the decision H_0 region of double sided Chi-square test is an interval $[\gamma_-, \gamma_+]$.*

As the function e^u/u is convex over $u \geq 0$ the H_0 decision region can be written in the form

$$\gamma_- \leq \frac{\hat{\sigma}_\mu^2}{\sigma_o^2} \leq \gamma_+,$$

where γ_- and γ_+ are selected to give $P_F = \alpha$.

A common choice of thresholds is ($\alpha \leq \frac{1}{2}$):

$$\begin{aligned}\gamma_- &= 1/n\ \chi_n^{-1}(\alpha/2)\\ \gamma_+ &= 1/n\ \chi_n^{-1}(1-\alpha/2)\end{aligned}$$

which gives equal area ($\alpha/2$) to the upper and lower tails of the χ_n distribution corresponding to a total FA probability $P_F = \alpha$ (see Fig. 132).

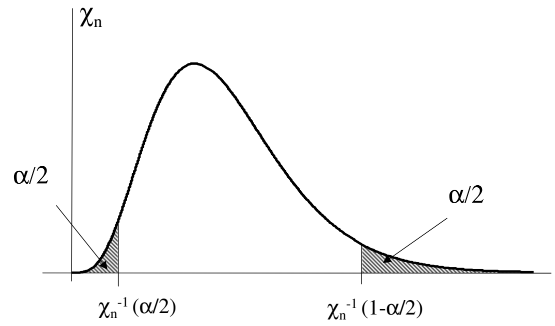


Figure 132: *Quantiles of Chi-square specify the thresholds γ_- and γ_+ for double sided test of variance.*

Power of double sided GLRT of variance:

Assume that the true value of $\sigma^2 > \sigma_o^2$ under H_1 is $\sigma^2 = \sigma_1^2$. Then

$$\begin{aligned}\beta = 1 - P(n\gamma_- \leq \frac{n\hat{\sigma}_\mu^2}{\sigma_o^2} \leq n\gamma_+ | H_1)\\ &= 1 - P(n\gamma_- \leq \underbrace{\frac{n\hat{\sigma}_\mu^2}{\sigma_1^2}}_{\chi_n \text{ under } H_1} \left(\frac{\sigma_1^2}{\sigma_o^2}\right) \leq n\gamma_+ | H_1)\\ &= 1 - \chi_n\left(n\gamma_+ \frac{\sigma_o^2}{\sigma_1^2}\right) + \chi_n\left(n\gamma_- \frac{\sigma_o^2}{\sigma_1^2}\right).\end{aligned}$$

9.4 TESTS ON VARIANCE: UNKNOWN MEAN

CASE I: $H_0 : \sigma^2 = \sigma_o^2,\ \mu \in \mathbb{R}$, $H_1 : \sigma^2 > \sigma_o^2,\ \mu \in \mathbb{R}$

CASE II: $H_0 : \sigma^2 < \sigma_o^2,\ \mu \in \mathbb{R}$, $H_1 : \sigma^2 > \sigma_o^2,\ \mu \in \mathbb{R}$

CASE III: $H_0 : \sigma^2 = \sigma_o^2,\ \mu \in \mathbb{R}$, $H_1 : \sigma^2 \neq \sigma_o^2\ \mu \in \mathbb{R}$

9.4.1 CASE I: $H_0 : \sigma^2 = \sigma_o^2,\ H_1 : \sigma^2 > \sigma_o^2$

We now have

$$\begin{aligned}\Lambda_{\mathrm{GLR}} &= \frac{\max_{\sigma^2>\sigma_o^2,\,\mu} f(\underline{x};\mu,\sigma^2)}{\max_\mu f(\underline{x};\mu,\sigma_o^2)} \\ &= \begin{cases} \frac{f(\underline{x};\,\overline{x},\,\hat{\sigma}^2)}{f(\underline{x};\,\overline{x},\sigma_o^2)}, & \hat{\sigma}^2 > \sigma_o^2 \\ 1, & \hat{\sigma}^2 \leq \sigma_o^2 \end{cases}\end{aligned}$$

where now

$$\hat{\sigma}^2 = n^{-1}\sum_{i=1}^n (x_i - \overline{x})^2$$

This is identical to the case of known μ with μ replaced by $\overline{x}$.

Hence, referring to work done for that case, we immediately obtain the single sided GLRT

$$T(\underline{x}) = \frac{n\hat{\sigma}^2}{\sigma_o^2} = \frac{(n-1)\mathsf{s}^2}{\sigma_o^2} \mathop{\gtrless}_{H_0}^{H_1} \gamma$$

PERFORMANCE

Under H_0,

* $T(\underline{X})$ is a Chi-square r.v. with $n-1$ d.f. and thus

$$\gamma = \chi_{n-1}^{-1}(1-\alpha)$$

Under H_1,

* $T(\underline{X})$ is Chi-square with $n-1$ d.f. (scaled by σ^2/σ_o) so

$$\begin{aligned}\beta &= 1 - \chi_{n-1}^{-1}(\gamma\ \sigma_o^2/\sigma^2) \\ &= 1 - \chi_{n-1}(\chi_{n-1}^{-1}(1-\alpha)\ \sigma_o^2/\sigma^2)\end{aligned}$$

9.4.2 CASE II: $H_0 : \sigma^2 < \sigma_o^2,\ \mu \in \mathbb{R}$, $H_1 : \sigma^2 > \sigma_o^2,\ \mu \in \mathbb{R}$

GLRT is dentical to Case I.

9.4.3 CASE III: $H_0 : \sigma^2 = \sigma_o^2,\ \mu \in \mathbb{R}$, $H_1 : \sigma^2 \neq \sigma_o^2\ \mu \in \mathbb{R}$

The derivation of the GLRT is completely analogous to Case III for known mean.

The H_0 decision region is identical to before except that sample variance replaces $\hat{\sigma}_\mu^2$, and the test statistic now has only $n-1$ d.f.

$$\chi_{n-1}(\alpha/2) \ \leq \ \frac{(n-1)\mathsf{s}^2}{\sigma_o^2} \ \leq \ \chi_{n-1}(1-\alpha/2)$$

The power function is identical to previous Case III for known μ except that χ_n CDF is replaced by χ_{n-1} CDF.

9.5 TESTS ON EQUALITY OF MEANS: UNKNOWN COMMON VARIANCE

Two i.i.d. independent samples

$\underline{X} = [X_1, \ldots, X_{n_1}]^T$, $X_i \ \sim \ \mathcal{N}(\mu_x, \sigma^2)$

$\underline{Y} = [Y_1, \ldots, Y_{n_2}]^T$, $Y_i \ \sim \ \mathcal{N}(\mu_y, \sigma^2)$

Case I: $H_0 : \mu_x = \mu_y,\ \sigma^2 > 0$, $H_1 : \mu_x \neq \mu_y,\ \sigma^2 > 0$

Case II: $H_0 : \mu_x \leq \mu_y,\ \sigma^2 > 0$, $H_1 : \mu_x > \mu_y,\ \sigma^2 > 0$

$\underline{X}, \underline{Y}$ have the joint density

$$\begin{aligned} f(\underline{x}, \underline{y};\ \mu_x, \mu_y, \sigma_x, \sigma_y) \ &= \ \left(\frac{1}{2\pi\sigma_x^2}\right)^{n_1/2} \left(\frac{1}{2\pi\sigma_y^2}\right)^{n_2/2} \\ &\cdot \exp\left(-\frac{1}{2\sigma_x^2}\ \sum_{i=1}^n (y_i - \mu_x)^2 - \frac{1}{2\sigma_y^2}\ \sum_{i=1}^n (y_i - \mu_y)^2\right) \end{aligned}$$

where $n = n_1 + n_2$.

9.5.1 CASE I: $H_0 : \mu_x = \mu_y,\ \sigma^2 > 0$, $H_1 : \mu_x \neq \mu_y,\ \sigma^2 > 0$

This is the case where X and Y have identical variances but possibly different means. The GLR statistic is given by

$$\Lambda_{\rm GLR} \ = \ \frac{\max_{\mu_x \neq \mu_y, \sigma^2 > 0} f(\underline{x}, \underline{y}; \mu_x, \mu_y, \sigma^2)}{\max_{\mu, \sigma^2 > 0} f(\underline{x}, \underline{y}; \mu, \mu, \sigma^2)} \tag{121}$$

$$= \ \frac{\max_{\mu_x, \mu_y, \sigma^2 > 0} f(\underline{x}, \underline{y}; \mu_x, \mu_y, \sigma^2)}{\max_{\mu, \sigma^2 > 0} f(\underline{x}, \underline{y}; \mu, \mu, \sigma^2)}. \tag{122}$$

This can be simplified. To do this note that the MLE for the case $\mu_x \neq \mu_y$ is

$$\hat{\mu}_x \ = \ \overline{x} = n_1^{-1} \sum_{i=1}^{n_1} x_i$$

$$\begin{aligned}
\hat{\mu}_y &= \overline{x} = n_2^{-1}\sum_{i=1}^{n_2} y_i \\
\hat{\sigma}_1^2 &= n^{-1}\sum_{i=1}^{n_1}(x_i-\overline{x})^2 + n^{-1}\sum_{i=1}^{n_2}(y_i-\overline{y})^2 \\
&= \frac{n_1}{n}\hat{\sigma}_x^2 + \frac{n_2}{n}\hat{\sigma}_y^2,
\end{aligned}$$

while the MLE for case $\mu_x = \mu_y = \mu$ is

$$\begin{aligned}
\hat{\mu} &= n^{-1}\sum_{i=1}^{n_1} x_i + n^{-1}\sum_{i=1}^{n_2} y_i = \frac{n_1}{n}\hat{\mu}_x + \frac{n_2}{n}\hat{\mu}_y \\
\hat{\sigma}_0^2 &= n^{-1}\sum_{i=1}^{n_1}(x_i-\hat{\mu})^2 + n^{-1}\sum_{i=1}^{n_2}(y_i-\hat{\mu})^2 \\
&= \hat{\sigma}_1^2 + \frac{n_1}{n}(\hat{\mu}-\overline{x})^2 + \frac{n_2}{n}(\hat{\mu}-\overline{y})^2
\end{aligned}$$

Plugging these two MLE's into the numerator and denominator of the LR statistic (122), we obtain after some simple algebra

$$\begin{aligned}
\Lambda_{\mathrm{GLR}} &= \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}\right)^{n/2} c \\
&= \left(\frac{\hat{\sigma}_1^2 + \frac{n_1}{n}(\hat{\mu}-\overline{x})^2 + \frac{n_2}{n}(\hat{\mu}-\overline{y})^2}{\hat{\sigma}_1^2}\right)^{n/2} c.
\end{aligned}$$

Thus one form of GLRT test is

$$\frac{\frac{n_1}{n}(\hat{\mu}-\overline{x})^2 + \frac{n_2}{n}(\hat{\mu}-\overline{y})^2}{\frac{n_1}{n}\hat{\sigma}_x^2 + \frac{n_2}{n}\hat{\sigma}_y^2} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma.$$

To reduce this to a well known test statistic we use the identities

$$\hat{\mu}-\overline{x} = \frac{n_1}{n}(\overline{y}-\overline{x}),$$

and

$$\hat{\mu}-\overline{y} = -\frac{n_2}{n}(\overline{y}-\overline{x}),$$

to obtain final form of GLRT (shown in Fig. 133)

$$T(\underline{x},\underline{y}) = \frac{|(\overline{y}-\overline{x})|}{\mathsf{s}_2\sqrt{\frac{n_1 n_2}{n}}} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma, \tag{123}$$

where we have defined the pooled sample variance

$$\mathsf{s}_2^2 = \frac{1}{n-2}\left(\sum_{i=1}^{n_1}(x_i-\hat{\mu})^2 + \sum_{i=1}^{n_2}(y_i-\hat{\mu})^2\right)$$

The test (123) is the well known *unpaired t-test.*

PERFORMANCE OF UNPAIRED T-TEST

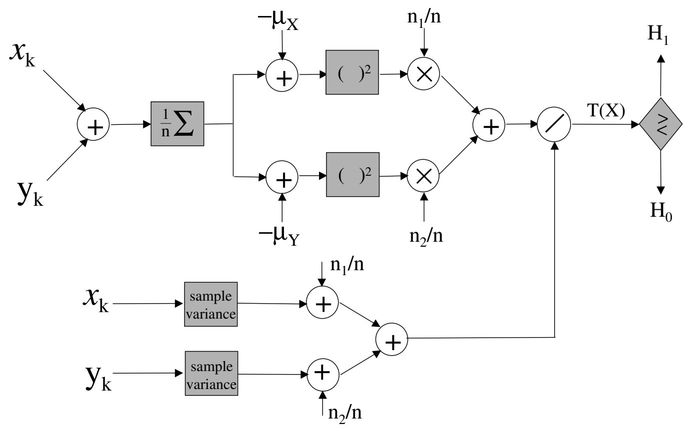


Figure 133: *Block diagram of test of equality of means of two populations.*

Under H_0:

$$\overline{Y_i} - \overline{X_i} = \mathcal{N}(0, \sigma^2) \cdot \sqrt{(1/n_1 + 1/n_2)},$$

and the test statistic is of the form of the magnitude of a Student-t random variable with $n_1 + n_2 - 2 = n - 2$ d.f:

$$T(\underline{X}, \underline{Y}) = \left| \frac{\mathcal{N}(0,1)}{\sqrt{\chi_{n-2}/(n-2)}} \right|.$$

Setting the GLRT threshold is now straightforward

$$\alpha = P_0(-\gamma < \mathcal{T}_{n-2} \leq \gamma)$$

and we conclude that $\gamma = \mathcal{T}_{n-2}^{-1}(1 - \alpha/2)$. This yields the level α unpaired-t test

$$\frac{\left|\sqrt{\frac{n_1 n_2}{n}}(\overline{y} - \overline{x})\right|}{\mathsf{s}_2} \overset{H_1}{\underset{H_0}{\gtrless}} \mathcal{T}_{n-2}^{-1}(1 - \alpha/2).$$

Under H_1 the test statistic equivalent to the magnitude of a non-central Student-t random variable with $n - 2$ d.f and non-centrality $d = \sqrt{n_1 n_2 / n}|\mu_y - \mu_x|/\sigma$.

9.5.2 CASE II: $H_0 : \mu_y \leq \mu_x,\ \sigma^2 > 0$, $H_1 : \mu_y > \mu_x,\ \sigma^2 > 0$

In an analogous manner to before we find that the GLRT reduces to the one sided t-test

$$\frac{\frac{\sqrt{n_1 n_2}}{n}(\overline{y} - \overline{x})}{\mathsf{s}_2} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma = \mathcal{T}_{n-2}^{-1}(1 - \alpha).$$

9.6 TESTS ON EQUALITY OF VARIANCES

Two i.i.d. independent samples

* $\underline{X} = [X_1, \ldots, X_{n_1}]^T$, $X_i \sim \mathcal{N}(\mu_x, \sigma_x^2)$

* $\underline{Y} = [Y_1, \ldots, Y_{n_2}]^T$, $Y_i \sim \mathcal{N}(\mu_y, \sigma_x^2)$

* μ_x, μ_y unknown

Case I: $H_0 : \sigma_x^2 = \sigma_y^2$, $H_1 : \sigma_x^2 \neq \sigma_y^2$

Case II: $H_0 : \sigma_x^2 = \sigma_y^2$, $H_1 : \sigma_x^2 > \sigma_y^2$

9.6.1 CASE I: $H_0 : \sigma_x^2 = \sigma_y^2$, $H_1 : \sigma_x^2 \neq \sigma_y^2$

The GLRT for testing equality of variances against the double sided alternative is

$$\begin{aligned}
\Lambda_{\rm GLR} &= \frac{\max_{\sigma_x^2 \neq \sigma_y^2, \mu_x, \mu_y} f(\underline{x}, \underline{y}; \mu_x, \mu_y, \sigma_x^2, \sigma_y^2)}{\max_{\sigma_x^2 = \sigma_y^2, \mu_x, \mu_y} f(\underline{x}, \underline{y}; \mu, \mu, \sigma_x^2, \sigma_y^2)} && (124) \\
&= \frac{f(\underline{x}, \underline{y}; \overline{x}, \overline{y}, \hat{\sigma}_x^2, \hat{\sigma}_y^2)}{f(\underline{x}, \underline{y}; \overline{x}, \overline{y}, \hat{\sigma}^2, \hat{\sigma}^2)}, && (125)
\end{aligned}$$

where we have defined the pooled variance estimate $\hat{\sigma}^2$ as

$$\hat{\sigma}^2 = \frac{n_1}{n}\hat{\sigma}_x^2 + \frac{n_2}{n}\hat{\sigma}_y^2.$$

The expression (125) is easily shown to reduce to

$$\Lambda_{\rm GLR} = \sqrt{\frac{(\hat{\sigma}^2)^{n_1+n_2}}{(\hat{\sigma}_x^2)^{n_1}\ (\hat{\sigma}_y^2)^{n_2}}} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \eta'.$$

Thus we obtain the equivalent GLRT test of equality of variances:

$$\sqrt{\underbrace{\left(1 + \overbrace{\frac{n_2 \hat{\sigma}_y^2}{n_1 \hat{\sigma}_x^2}}^{u}\right)^{n_1} \cdot \left(1 + \overbrace{\frac{n_1 \hat{\sigma}_x^2}{n_2 \hat{\sigma}_y^2}}^{1/u}\right)^{n_2}}_{g(u)}} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \eta. \qquad (126)$$

By investigating stationary points of the function $g(u) = (1+u)^{n_1}(1+1/u)^{n_2}$ it is easily established that $g(u)$ is convex and has a single minimum over the range $u \geq 0$. Specifically, note that (see Fig. 134)

$$g'(u) = \frac{n_1}{u^2}(1+u)^{n_1-1}(1+1/u)^{n_2-1}\ \left(u^2 + (1 - n_2/n_1)u - n_2/n_1\right)$$

has only one positive root which occurs at $u = n_2/n_1$. Hence the H_0 decision region for the GLRT is of the form

$$\gamma_- \leq \frac{n_2 \hat{\sigma}_y^2}{n_1 \hat{\sigma}_x^2} \leq \gamma_+,$$

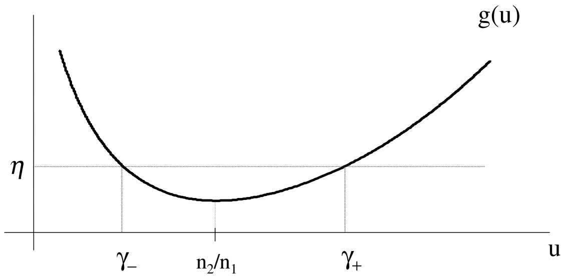


Figure 134: *The double sided test statistic is of the form $g(u)$ which is convex over $u \geq 0$ with minimum at $u = n_2/n_1$.*

which is equivalent to a Fisher F-test (see Fig. 135)

$$\gamma_-' \leq \frac{\mathsf{s}_y^2}{\mathsf{s}_x^2} \leq \gamma_+'.$$

The thresholds γ_- and γ_+ can be set according to

$$\begin{aligned} \gamma_-' &= \mathcal{F}^{-1}_{n_1-1,n_2-1}(\alpha/2) \\ \gamma_+' &= \mathcal{F}^{-1}_{n_1-1,n_2-1}(1-\alpha/2). \end{aligned}$$

9.6.2 CASE II: $H_0 : \sigma_x^2 = \sigma_y^2$, $H_1 : \sigma_y^2 > \sigma_x^2$

The GLRT for testing equality of variances against the single sided alternative is

$$\begin{aligned} \Lambda_{\text{GLR}} &= \frac{\max_{\sigma_y^2 > \sigma_x^2, \mu_x, \mu_y} f(\underline{x}, \underline{y}; \mu_x, \mu_y, \sigma_x^2, \sigma_y^2)}{\max_{\sigma_y^2 = \sigma_x^2, \mu_x, \mu_y} f(\underline{x}, \underline{y}; \mu, \mu, \sigma_x^2, \sigma_y^2)} \\ &= \left\{ \begin{array}{cc} \sqrt{\frac{(\hat{\sigma}^2)^{n_1+n_2}}{(\hat{\sigma}_x^2)^{n_1}\,(\hat{\sigma}_y^2)^{n_2}}}, & \hat{\sigma}_y^2 > \hat{\sigma}_x^2 \\ 1, & \hat{\sigma}_y^2 = \hat{\sigma}_x^2 \end{array} \right. . \end{aligned}$$

From our study of the double sided case in Sec. 9.6.1 we know that the function $g(u)$ defined in (126) is convex with minimum at n_2/n_1. This implies that Λ_{GLR} is monotone increasing in $\hat{\sigma}_y^2/\hat{\sigma}_x^2$ over the range $\hat{\sigma}_y^2 > \hat{\sigma}_x^2$. Thus we obtain the GLRT as a single sided F-test

$$\frac{\mathsf{s}_y^2}{\mathsf{s}_x^2} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma,$$

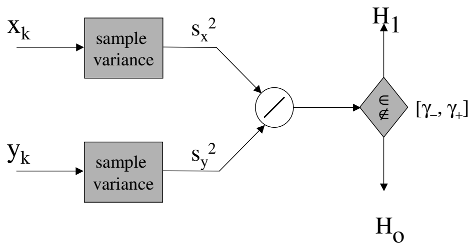


Figure 135: *Block diagram of test of equality of two variances.*

where

$$\gamma = \mathcal{F}^{-1}_{n_2-1,n_1-1}(1-\alpha).$$

9.7 TESTS ON CORRELATION

Assume that one has n i.i.d. pairs $\underline{Z}_i = [X_i, Y_i]$, $i = 1, \ldots, n$, of samples from a bivariate Gaussian density with unknown mean $\underline{mu} = [\mu_x, \mu_y]$ and unknown covariance

$$\mathbf{R} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix}.$$

The objective is to test whether or not the correlation between X_i and Y_i is zero. The sequence of i.i.d. vectors $\{\underline{Z}_i\}_{i=1}^n$ has the bivariate density

$$f(\underline{z};\ \underline{\mu}, \mathbf{R}) = \left(\frac{1}{2\pi\sqrt{|\det \mathbf{R}|}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n (\underline{z}_i - \underline{\mu})^T \mathbf{R}^{-1} (\underline{z}_i - \underline{\mu})\right).$$

Define the correlation coefficient ρ

$$\rho = \frac{\sigma_{xy}}{\sigma_x\, \sigma_y}.$$

As usual we consider testing the double sided and single sided hypotheses: Case I: $H_0 : \rho = 0$, $H_1 : \rho \neq 0$.

Case II: $H_0 : \rho = 0$, $H_1 : \rho > 0$.

9.7.1 CASE I: $H_0 : \rho = \rho_o$, $H_1 : \rho \neq \rho_o$

We will show in Sec. 12 that the maximum of the bivariate Gaussian p.d.f. $f(\underline{x}, \underline{y};\ \mu_x, \mu_y, \mathbf{R})$ over μ_x, μ_y and $\mathbf{R}$ is equal to

$$\max_{\mathbf{R},\mu_x,\mu_y} f(\underline{x}, \underline{y};\ \mu_x, \mu_y, \mathbf{R}) = \left(\frac{1}{(2\pi)^2 |\det \hat{\mathbf{R}}|}\right)^{n/2} e^{-n/2}$$

and that the maximum is attained by the joint ML estimates

$$\begin{aligned} \hat{\mu}_x &= \overline{X} \\ \hat{\mu}_y &= \overline{Y} \\ \hat{\mathbf{R}} &= \frac{1}{n}\sum_{i=1}^n \begin{bmatrix} x_i \\ y_i \end{bmatrix} [x_i, y_i] = \begin{bmatrix} \hat{\sigma}_x^2 & \hat{\sigma}_{xy} \\ \hat{\sigma}_{yx} & \hat{\sigma}_y^2 \end{bmatrix} \end{aligned}$$

where

$$\hat{\sigma}_{xy} = n^{-1}\sum_{i=1}^n (X_i Y_i - \overline{XY}),$$

and $\overline{XY} = n^{-1}\sum_{i=1}^n X_i Y_i$.

Using this we can easily find GLR statistic

$$\begin{aligned} \Lambda_{\text{GLR}} &= \frac{\max_{\mathbf{R},\mu_x,\mu_y} f(\underline{x}, \underline{y}; \mu_x, \mu_y, \mathbf{R})}{\max_{\text{diagonal } \mathbf{R},\mu_x,\mu_y} f(\underline{x}, \underline{y}; \mu_x, \mu_y, \mathbf{R})} \\ &= \left(\frac{\hat{\sigma}_x^2\hat{\sigma}_y^2}{\hat{\sigma}_x^2\hat{\sigma}_y^2 - \hat{\sigma}_{xy}^2}\right)^{n/2} \\ &= \left(\frac{1}{1-\hat{\rho}^2}\right)^{n/2}, \end{aligned}$$

where we have defined *sample correlation coefficient*

$$\hat{\rho} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x\ \hat{\sigma}_y}.$$

As Λ_{GLR} is monotonic increasing in $|\hat{\rho}|$ we have one simple form of the GLRT

$$|\hat{\rho}| \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma. \qquad (127)$$

An expression for the distribution under H_0 of $\hat{\rho}(\underline{X})$ can be derived [57] but it is not a standard well tabulated distribution.

A different form of the test is obtained by making a transformation on $\hat{\rho}$

$$g(u) = u^2/(1-u^2),$$

which is monotone increasing in $|u|$. Therefore, the GLRT (127) is equivalent to

$$\frac{|\hat{\rho}|}{\sqrt{1-\hat{\rho}^2}\sqrt{(n-2)}} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma.$$

The statistic $\hat{\rho}/\sqrt{1-\hat{\rho}^2}\sqrt{(n-2)}$ can be shown to follow the student-t distribution with $n-2$ d.f. [7]. Thus the level α threshold for this GLRT is

$$\gamma = \mathcal{T}_{n-2}^{-1}(1-\alpha/2).$$

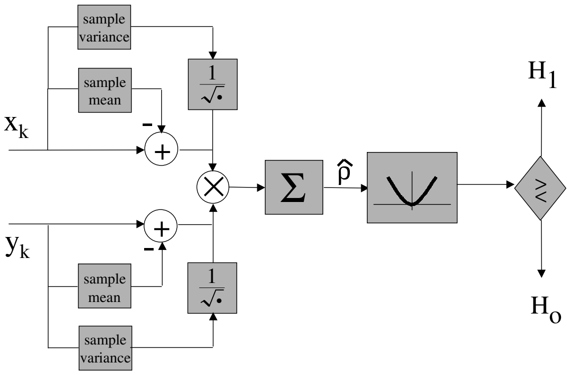


Figure 136: *Block diagram of double sided test of nonzero correlation between two populations.*

9.7.2 CASE II: $H_0 : \rho = 0$, $H_1 : \rho > 0$

By analogous methods as used to obtain the GLRT for one-sided tests on the variance in Sec. sec:GLRTscalarvar1, it can be shown that the GLRT for the one sided test of the correlation is of the form

$$\hat{\rho} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma,$$

or alternatively

$$\frac{(n-2)\,\hat{\rho}}{\sqrt{1-\hat{\rho}^2}} \underset{H_0}{\overset{H_1}{\gtrless}} \mathcal{T}_{n-2}^{-1}(1-\alpha).$$

9.8 BACKGROUND REFERENCES

The GLRT for i.i.d. scalar Gaussian measurements is well described in the statistics books by Mood, Graybill and Boes [48] and by Bickel and Docksum [7]. Coverage from a more applied engineering perspective is in [75].

9.9 EXERCISES

1. n i.i.d. realizations of a bivariate Gaussian random vector $\underline{z} = [z_1, z_2]^T$ are observed where the mean of $\underline{z}$ is $\underline{s}$ and the covariance is of the form:

$$\mathbf{R}_z = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \sigma^2, \qquad \text{Note :} \mathbf{R}_z^{-1} = \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix} \frac{1}{\sigma^2(1-\rho^2)}$$

where the component variances σ^2 and the correlation coefficient $\rho \in [-1, 1]$ are known.

(a) Derive the MP LRT (with threshold) of the simple hypotheses

$$\begin{aligned} H_0 &: \underline{s} = \underline{s}_0 \\ H_1 &: \underline{s} = \underline{s}_1. \end{aligned}$$

For $\underline{s}_0 = 0$ is your test UMP for $H_1 : \underline{s} \neq 0$? when $|\rho| > 0$? How about for $\rho = 0$?.

(b) Find the ROC of the MP LRT of part (a) and show that it is specified by the detectability index

$$d^2 = \frac{|E[T|H_1] - E[T|H_0]|^2}{\operatorname{var}(T|H_0)} = n(\underline{s}_1 - \underline{s}_0)^T \mathbf{R}_z^{-1} (\underline{s}_1 - \underline{s}_0)$$

where T is a test statistic linear in $\underline{z}$.

(c) Now assume that $\underline{s}_0$ and $\underline{s}_1$ satisfy the "power" constraints $\underline{s}_0^T \underline{s}_0 = \underline{s}_1^T \underline{s}_1 = 1$. For fixed ρ show that the optimal signal pair $\underline{s}_1, \underline{s}_0$ which maximizes d^2 must satisfy $\underline{s}_1 - \underline{s}_0 = c[1, 1]$ when $\rho < 0$ while it must satisfy $\underline{s}_1 - \underline{s}_0 = c[1, -1]$ when $\rho > 0$, where c is a constant ensuring the power constraint.

(d) Assuming the optimal signal pair derived in part (b), for what value of ρ is the detectability index the worst (smallest) and for what value is it the best? Does this make sense?

2. The test on equality of a pair of means in Sec. 9.5 required equal variances and led to the unpaired t-test of (123). Extend this to the case where the variances on each population may be unequal. This is a challenging problem.

End of chapter

10 STATISTICAL CONFIDENCE INTERVALS

In many cases an estimate of an unknown parameter does not suffice; one would also like to know something about the precision of the estimate. The estimation strategies discussed in Chapter 4 do not provide any help here. If one knew the true parameter θ, the detection strategies discussed in Chapter 7 could be used to specify precision of a parameter estimate $\hat{\theta}$ by testing the hypothesis that $\|\hat{\theta}-\theta\|$ is small. What one needs here is a different approach. Here we discuss the framework of statistical confidence intervals. In this framework, instead of seeking an estimate of the true parameter, called a point estimate, one seeks a tight interval that covers the true parameter with specified confidence level. It turns out that confidence intervals are closely related to tests of composite double sided hypotheses and we will exploit this connection in the presentation.

The specific topics in this chapter are:

OUTLINE

* Confidence intervals via pivots

* Confidence intervals and double sided tests

* Confidence interval for mean, variance, correlation

10.1 DEFINITION OF A CONFIDENCE INTERVAL

Let $\theta \in \Theta$ be an unknown scalar parameter and let $X \sim f(x;\theta)$ be an observed random variable, random vector or random process. As opposed to a point estimator $\hat{\theta}(X)$ which is a (random) point in the parameter space Θ, a confidence interval $[T_1(X), T_2(X)]$ is a (random) interval in parameter space. Confidence intervals are also called set or interval estimators.

OBJECTIVE: find two statistics $T_1 = T_1(X)$ and $T_2 = T_2(X)$, $T_1 < T_2$, which specify endpoints of a random interval

$$[T_1,\ T_2]$$

that contains θ with high probability.

CONSTRUCTION OF CONFIDENCE INTERVALS

1. Fix $\alpha \in [0,1]$

2. For all $\theta \in \Theta$ we require

$$P_\theta(T_1 \leq \theta \leq T_2) \geq 1-\alpha$$

The interval $[T_1, T_2]$ is called a $100(1-\alpha)\%$ confidence interval

Equivalently:

$\Rightarrow$ $1-\alpha$ is confidence level of statement "$\theta \in [T_1, T_2]$"

$\Rightarrow$ $[T_1, T_2]$ is a set estimate of θ

* $P_\theta(T_1 \leq \theta \leq T_2)$ is coverage probability of the confidence interval

* $1-\alpha$ is lower bound on coverage probability of a $100(1-\alpha)\%$ conf. interval

* $T_2 - T_1$ is the length of the confidence interval and, everything else being equal, we would like this to be as small as possible.

Figure 137: *Confidence interval covers θ with probability at least $1-\alpha$.*

* Sometimes a level α GLRT or LMPT of double sided hypotheses can be useful for finding confidence intervals.

10.2 CONFIDENCE ON MEAN: KNOWN VAR

Objective: find confidence interval on the mean μ based on i.i.d. Gaussian sample $\underline{X} = [X_1, \ldots, X_n]$ with known variance σ^2.

APPROACH 1: Use Tchebychev inequality:

$$P_\mu(|\overline{X_i} - \mu| \geq \epsilon) \leq \frac{\overbrace{E_\mu[(\overline{X_i} - \mu)^2]}^{\sigma^2/n}}{\epsilon^2}$$

or, setting $\epsilon = c\,\sigma/\sqrt{n}$

$$P_\mu(|\overline{X_i} - \mu| \geq c\,\sigma/\sqrt{n}) \leq \frac{1}{c^2}$$

or equivalently

$$P_\mu(|\overline{X_i} - \mu| \leq c\,\sigma/\sqrt{n}) \geq 1 - \frac{1}{c^2}$$

i.e.

$$P_\mu(\overline{X_i} - c\,\sigma/\sqrt{n} \leq \mu \leq \overline{X_i} + c\,\sigma/\sqrt{n}) \geq 1 - \frac{1}{c^2}$$

Finally take $c = 1/\sqrt{\alpha}$ to obtain $100(1-\alpha)\%$ confidence interval for μ

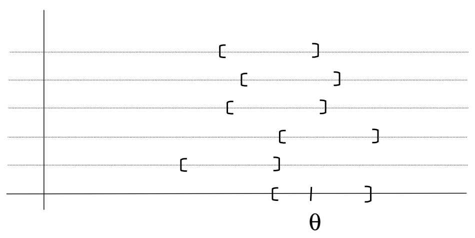


Figure 138: *In n-trials a $(1-\alpha)\%$ confidence interval $[T_1, T_2]$ covers θ approximately $(1-\alpha)n$ times (or more) for n large. Here shown is an* 85% *confidence interval.*

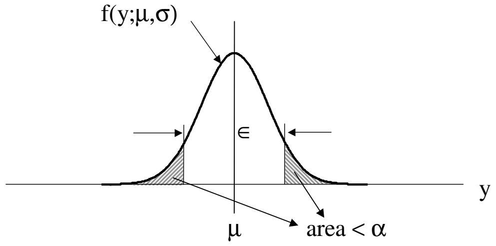


Figure 139: *Tchebychev inequality specifies an interval containing the mean with at least probability $1-\sigma^2/n\epsilon^2$. In figure $f(y;\mu,\sigma)$ denotes the density of the sample mean $Y = \overline{X_i}$.*

$$\left[\overline{X_i} - \frac{\sigma}{\sqrt{\alpha n}}, \ \overline{X_i} + \frac{\sigma}{\sqrt{\alpha n}}\right]$$

OBSERVATIONS:

* Tchebychev interval is symmetric about sample mean

* Size $2\frac{\sigma}{\sqrt{\alpha n}}$ of interval increases in σ^2/n

* There is a tradeoff between coverage probability $\geq 1-\alpha$ and small size

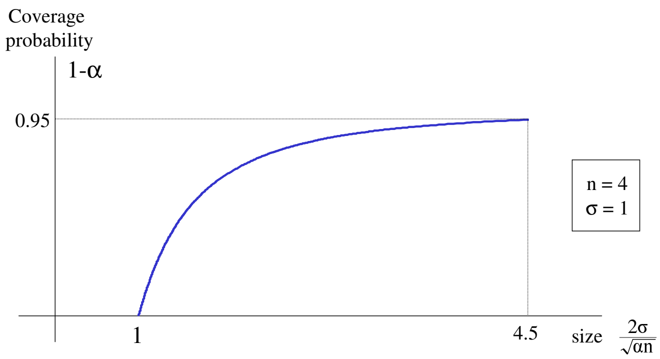


Figure 140: *Coverage vs. size of Tchebychev confidence interval for mean for a Gaussian sample having known variance.*

* Tchebychev interval is "distribution free"

* Actual coverage probability may be $\gg$ desired confidence $1-\alpha$

* Tchebychev intervals are usually excessively large

APPROACH 2: Find exact confidence interval by finding a *pivot*

Recall problem of testing double sided hypotheses for i.i.d. Gaussian r.v.s having known variance

$$\begin{aligned} H_0 : \mu &= \mu_o \\ H_1 : \mu &\neq \mu_o \end{aligned}$$

we found level α GLRT

$$|Q(\underline{x}, \mu_o)| = \frac{\sqrt{n}|\overline{x_i} - \mu_o|}{\sigma} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma = \mathcal{N}^{-1}(1-\alpha/2)$$

where we have defined the function

$$Q(\underline{x}, \mu_o) = \frac{\sqrt{n}(\mu_o - \overline{x_i})}{\sigma}.$$

Note when $\mu = \mu_o$:

1) $Q(\underline{X}, \mu_o)$ satisfies the following proeprties

– it is a monotone function of μ_o

– it has a probability distribution independent of μ_o (and σ):

$\Rightarrow$ such a function $Q(\underline{X}, \mu_o)$ is called a PIVOT. it has also been called a "root" [6].

2) By design of the threshold $\gamma = \mathcal{N}^{-1}(1-\alpha/2)$ the false alarm probability of the test is

$$P_{\mu_o}\left(|Q(\underline{X}, \mu_o)| > \gamma\right) = \alpha$$

for arbitrary μ_o.

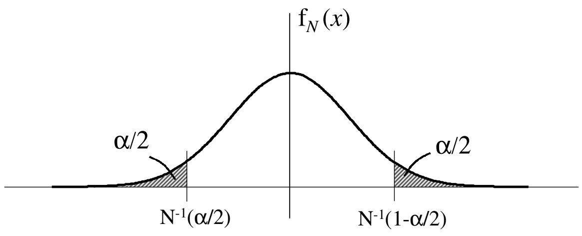


Figure 141: *The false alarm level setting for double sided test mean gives an exact* $1-\alpha$ *confidence interval.*

As P_F is independent of μ_o, μ_o is just a dummy variable which can be replaced by the generic μ and

$$P_\mu\left(-\gamma \le Q(\underline{X}, \mu) \le \gamma\right) = 1 - \alpha.$$

As $Q(\underline{X}, \mu) = \sqrt{n}(\overline{x_i} - \mu)/\sigma$ is monotone in μ the following inequalities on Q: $-\gamma \le Q(\underline{X}, \mu) \le \gamma$, are equivalent to the following inequalities on μ: $\overline{X_i} - \frac{\sigma}{\sqrt{n}}\gamma \le \mu \le \overline{X_i} + \frac{\sigma}{\sqrt{n}}\gamma$. Thus we have found the following EXACT $100(1-\alpha)\%$ conf. interval for μ

$$\left[\overline{X_i} - \frac{\sigma}{\sqrt{n}}\mathcal{N}^{-1}(1-\alpha/2),\ \overline{X_i} + \frac{\sigma}{\sqrt{n}}\mathcal{N}^{-1}(1-\alpha/2)\right]$$

OBSERVATIONS

1. By the central limit theorem this interval is accurate for large n even for non-Gaussian case

$$\sqrt{n}(\overline{X_i}-\mu)/\sigma \;\rightarrow\; \mathcal{N}(0,1), \quad (i.d.)$$

2. Exact interval is symmetric about $\overline{X_i}$
3. Exact interval is significantly smaller than Tchebychev

$$[T_2 - T_1]_{Tchby} = 2\frac{\sigma}{\sqrt{\alpha n}} > 2\frac{\sigma}{\sqrt{n}}\,\mathcal{N}^{-1}(1-\alpha/2) = [T_2 - T_1]_{Exact}$$

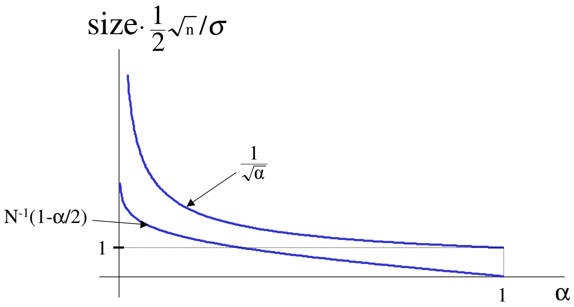


Figure 142: *Size vs. (1-confidence level) for exact and Tchebychev intervals.*

10.3 CONFIDENCE ON MEAN: UNKNOWN VAR

Objective: find conf. inerval on the mean μ based on i.i.d. Gaussian sample $\underline{X} = [X_1, \ldots, X_n]$ with unknown variance σ^2.

APPROACH: Exact confidence interval via pivot

Solution: Motivated by previous example we go back to the double sided hypothesis GLRT for unknown variance

$$\begin{aligned} H_0 : \mu &= \mu_o, \quad \sigma^2 > 0 \\ H_1 : \mu &\neq \mu_o, \quad \sigma^2 > 0 \end{aligned}$$

we found level α t-test for Gaussian x_i's:

$$|Q(\underline{x}, \mu_o)| = \frac{\sqrt{n}|\overline{x_i} - \mu_o|}{\mathsf{s}} \;\overset{H_1}{\underset{H_0}{\gtrless}}\; \gamma = \mathcal{T}_{n-1}^{-1}(1-\alpha/2)$$

Therefore, an exact $(1-\alpha)\%$ confidence interval for μ is

$$\left[\overline{X_i} - \frac{\mathsf{s}}{\sqrt{n}}\,\mathcal{T}_n^{-1}(1-\alpha/2),\ \overline{X_i} + \frac{\mathsf{s}}{\sqrt{n}}\,\mathcal{T}_{n-1}^{-1}(1-\alpha/2)\right]$$

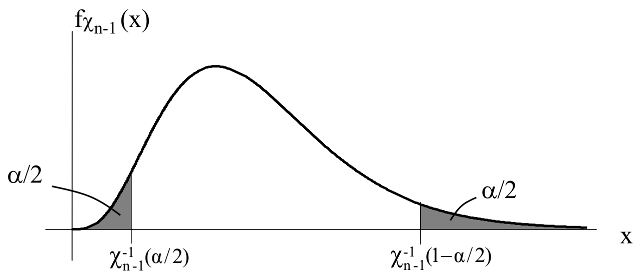


Figure 143: *Exact confidence interval for mean for unknown variance in a Gaussian sample is given by quantiles of student-t distribution with $n-1$ d.f.*

10.4 CONFIDENCE ON VARIANCE

Objective: find conf. interval on variance σ^2 based on i.i.d. Gaussian sample $\underline{X} = [X_1, \ldots, X_n]$ with unknown mean μ.

Solution: Recall the double sided hypothesis GLRT for variance

$$H_0 : \sigma^2 = \sigma_o^2,$$

$$H_1 : \sigma^2 \neq \sigma_o^2,$$

In a previous chapter we found a level α Chi-square test for Gaussian X_i's in terms of the sample variance s^2:

$$\chi_{n-1}^{-1}(\alpha/2) \leq \frac{(n-1)\mathsf{s}^2}{\sigma_o^2} \leq \chi_{n-1}^{-1}(1-\alpha/2)$$

Therefore, an exact $100(1-\alpha)\%$ confidence interval for σ^2 is

$$\left[\frac{(n-1)\mathsf{s}^2}{\chi_{n-1}^{-1}(1-\alpha/2)},\ \frac{(n-1)\mathsf{s}^2}{\chi_{n-1}^{-1}(\alpha/2)}\right]$$

Note: confidence interval for variance is not symmetric about s^2

10.5 CONFIDENCE ON DIFFERENCE OF TWO MEANS

Objective: find conf. interval on the difference $\Delta = \mu_x - \mu_y$ of means in two i.i.d. Gaussian samples $\underline{X} = [X_1, \ldots, X_{n_1}]$, $\underline{Y} = [Y_1, \ldots, Y_{n_2}]$

Solution: consider the hypotheses

$$\begin{aligned} H_0 : \Delta &= \Delta_o \\ H_1 : \Delta &\neq \Delta_o \end{aligned}$$

A GLRT of level α would give a confidence interval for Δ similiarly to before.

Recall: the t test

$$\frac{\sqrt{\frac{n_1 n_2}{n}}|\overline{x_i} - \overline{y_i}|}{\mathsf{s}_2} \overset{H_1}{\underset{H_0}{\gtrless}} \mathcal{T}^{-1}_{n-2}(1-\alpha/2)$$

was previously derived for the double sided hypotheses

$$\begin{aligned} H_0' : \mu_x &= \mu_y \\ H_1' : \mu_x &\neq \mu_y \end{aligned}$$

There is a difficulty, however, since Δ does not appear any where in the test statistic. In particular

* $\overline{X_i} - \overline{Y_i}$ has mean Δ under $H_0 : \Delta = \Delta_o$

* therefore distribution of t-test statistic above depends on Δ and is not a pivot

However, as $\overline{X_i} - \overline{Y_i} - \Delta$ has mean zero under H_0 and same variance as $\overline{X_i} - \overline{Y_i}$, we can immediately identify the following pivot

$$\frac{\sqrt{\frac{n_1 n_2}{n}}(\overline{X_i} - \overline{Y_i} - \Delta)}{\mathsf{s}_2} \sim \mathcal{T}_{n-2}$$

Thus, the left and right endpoints of a $100(1-\alpha)\%$ conf. interval on Δ are given by

$$\overline{X_i} - \overline{Y_i} \mp \sqrt{\frac{n}{n_1 n_2}}\, \mathsf{s}_2\, \mathcal{T}^{-1}_{n-2}(1-\alpha/2)$$

10.6 CONFIDENCE ON RATIO OF TWO VARIANCES

Objective: find conf. interval on the ratio

$$c = \sigma_x^2/\sigma_y^2$$

of variances in two i.i.d. Gaussian samples

* $\underline{X} = [X_1, \ldots, X_{n_1}]$, $\underline{Y} = [Y_1, \ldots, Y_{n_2}]$

Solution: Recall that the GLRT for double sided $H_1 : \sigma_x^2 \neq \sigma_y^2$ was F-test

$$\mathcal{F}^{-1}_{n_1-1,n_2-1}(\alpha/2) \leq \frac{\mathsf{s}_x^2}{\mathsf{s}_y^2} \leq \mathcal{F}^{-1}_{n_1-1,n_2-1}(1-\alpha/2)$$

Difficulty: distribution of test statistic depends on $c = \sigma_x^2/\sigma_y^2$

However, as

$$\frac{1}{c}\frac{\mathsf{s}_X^2}{\mathsf{s}_Y^2} = \frac{\mathsf{s}_X^2/\sigma_X^2}{\mathsf{s}_Y^2/\sigma_Y^2} \sim \mathcal{F}_{n_1-1,n_2-1}$$

we have identified a pivot.

Therefore, a $(1-\alpha)\%$ conf. interval on variance ratio $c = \sigma_x^2/\sigma_y^2$ is given by

$$[T_1, T_2] = \left[\frac{\mathsf{s}_X^2}{\mathsf{s}_Y^2\, \mathcal{F}^{-1}_{n_1-1,n_2-1}(1-\alpha/2)}, \frac{\mathsf{s}_X^2}{\mathsf{s}_Y^2\, \mathcal{F}^{-1}_{n_1-1,n_2-1}(\alpha/2)}\right]$$

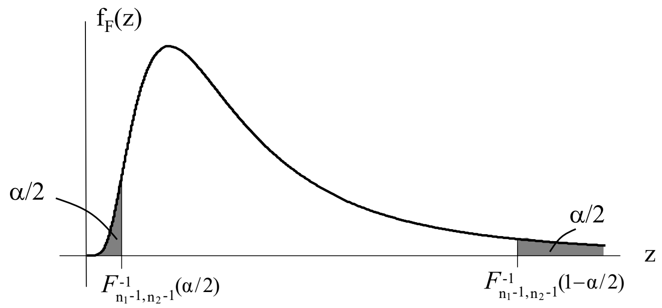


Figure 144: *Confidence interval on variance ratio in a pair of Gaussian samples depends on quantiles of F-distribution* $\mathcal{F}_{n_1-1,n_2-1}$

10.7 CONFIDENCE ON CORRELATION COEFFICIENT

Objective: find conf. interval on correlation coefficient ρ between two i.i.d. Gaussian samples $\underline{X} = [X_1, \ldots, X_n]$, $\underline{Y} = [Y_1, \ldots, Y_n]$ with unknown means and variances.

NOTE: not obvious how to obtain pivot from previously derived GLRT test statistic for testing $H_1 : \rho \neq 0$.

Solution: Fisher Transformation.

Let $\hat{\rho}$ be sample correlation coefficient. Then

$$\upsilon = \tfrac{1}{2} \ln \left(\frac{1+\hat{\rho}}{1-\hat{\rho}} \right) = \tanh^{-1}(\hat{\rho})$$

has an asymptotic normal dsn with

$$E_\theta[\upsilon] = \tanh^{-1}(\rho)$$
$$\mathrm{var}_\theta(\upsilon) = \frac{1}{n-3}$$

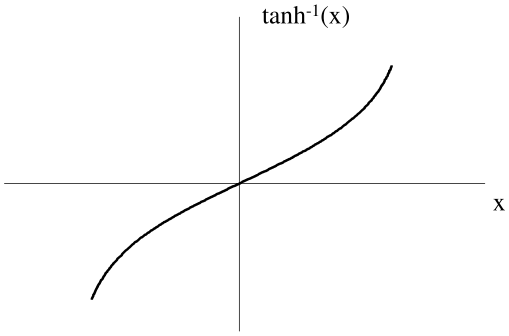


Figure 145: *Inverse tanh function is monotone increasing.*

Hence we have a pivot

$$Q(\underline{X}, \rho) = \frac{\tanh^{-1}(\hat{\rho}) - \tanh^{-1}(\rho)}{1/\sqrt{n-3}} \sim \mathcal{N}(0,1)$$

This gives $100(1-\alpha)\%$ conf. interval on $\tanh^{-1}(\rho)$

$$\left[\upsilon - \frac{1}{\sqrt{n-3}}\, \mathcal{N}^{-1}(1-\alpha/2),\ \upsilon + \frac{1}{\sqrt{n-3}}\, \mathcal{N}^{-1}(1-\alpha/2) \right]$$

Since $\tanh^{-1}(\cdot)$ is monotone, the left and right endpoints T_1, T_2 of $(1-\alpha)\%$ conf. interval $[T_1, T_2]$ on ρ are

$$\tanh \left(\upsilon \mp \frac{1}{\sqrt{n-3}}\, \mathcal{N}^{-1}(1-\alpha/2) \right)$$

OBSERVATIONS:

1. Conf. interval is not symmetric

2. Conf. interval prescribes a level α test of double sided hypotheses

$$\begin{aligned} H_0 : \rho &= \rho_o, \\ H_1 : \rho &\neq \rho_o \end{aligned}$$

which is

$$\phi(\underline{x}) = \begin{cases} 1, & \rho_o \notin [T_1, T_2] \\ 0, & \rho_o \in [T_1, T_2] \end{cases}$$

Indeed:

$$\begin{aligned} E_0[\phi] &= 1 - P_{\rho_o}(T_1 \leq \rho_o \leq T_2) \\ &= 1 - (1-\alpha) = \alpha \end{aligned}$$

10.8 BACKGROUND REFERENCES

There are other ways to construct confidence intervals besides exploiting the double sided GLRT relationship. As in previous chapters we refer the reader to the two excellent books by Mood, Graybill and Boes [48] and by Bickel and Docksum [7] for more general discussion of confidence intervals. The book by Hjorth [26] covers the theory and practice of bootstrap confidence intervals, a powerful nonparametric but computationally intensive approach to the interval estimation problem. A generalization of the theory of confidence intervals is the theory of confidence regions, which is briefly presented in Sec. 12.6 after we discuss the GLRT of multivariate double sided hypotheses.

10.9 EXERCISES

10.1 Let $\{X_i\}_{i=1}^n$ be an i.i.d. sample with marginal p.d.f. $f(x;\theta) = \theta e^{-\theta x}$, $x > 0$.

(a) Show that the maximum likelihood estimator (MLE) for θ is $\frac{1}{\bar{X}}$, where $\bar{X}$ is the sample mean (you should have previously derived this in hwk 2).

(b) Show that the CR bound $I^{-1}(\theta)$ for θ given $\{X_i\}_{i=1}^n$ is of the form: $I^{-1}(\theta) = \frac{\theta^2}{n}$.

(c) Now, using the fact that for large n the MLE is distributed as approximately $\mathcal{N}(\theta, I^{-1}(\theta))$, show that $\left(\frac{1/\bar{X}}{1+Z(1-\frac{\alpha}{2})/\sqrt{n}}, \frac{1/\bar{X}}{1-Z(1-\frac{\alpha}{2})/\sqrt{n}}\right)$ is a $(1-\alpha)\cdot 100\%$ confidence interval for θ, where $Z(p) = \mathcal{N}^{-1}(p) = \left\{x : \int_{-\infty}^{x} e^{-\frac{1}{2}x^2} dx/\sqrt{2\pi} = p\right\}$ is the $p-th$ quantile of $\mathcal{N}(0,1)$.

10.2 Let $\underline{\mathbf{X}} = [\mathbf{X}_1, \ldots, \mathbf{X}_n]^T$ be a Gaussian random vector with mean $\underline{\mu} = [\mu_1, \ldots, \mu_n]^T$ and covariance matrix R_X.

(a) Show that the distribution of $\mathbf{W} \stackrel{\text{def}}{=} (\underline{\mathbf{X}} - \underline{\mu})^T R_X^{-1} (\underline{\mathbf{X}} - \underline{\mu})$ is Chi-Square with n degrees of freedom (Hint: use square root factor $R_X^{\frac{1}{2}}$ to represent $(\underline{\mathbf{X}} - \underline{\mu})$ in terms of a vector of uncorrelated standard Gaussian variates).

(b) Since $\mathbf{W}$ has a distribution which is independent of $\underline{\mu}$ and R_X, $\mathbf{W}$ is similar to a pivot for scalar μ which can be used to generate confidence regions on $\underline{\mu}$. Assume $n = 2$ and let R_X be a fixed and known diagonal matrix with eigenvalues λ_1 and λ_2. Show that $\mathcal{R} \stackrel{\text{def}}{=} \{\underline{\mu} : (\underline{\mathbf{X}} - \underline{\mu})^T R_X^{-1} (\underline{\mathbf{X}} - \underline{\mu}) \leq -2\ln\alpha\}$ is a $100(1-\alpha)\%$ confidence region for the vector $\underline{\mu}$ in the sense that: $P(\underline{\mu} \in \mathcal{R}) = P(W \leq -2\ln\alpha) = 1-\alpha$. Draw a concise picture of this confidence region for the case $\underline{\mu} \in \mathbb{R}^2$. Label and identify all quantities in your picture. What happens to the confidence region as $\lambda_1 \to 0$? Does this make sense?

10.3 *This exercise establishes that a pivot always exists when the marginal CDF is strictly increasing.* Let $\{X_i\}_{i=1}^n$ be an i.i.d. sample with marginal p.d.f. $f(x;\theta)$ and a CDF $F(x;\theta)$ which is strictly increasing: $F(x+\Delta;\theta+\delta) > F(x;\theta)$, $\Delta, \delta > 0$.

(a) Show that the random variable $F(\mathbf{X}_i;\theta)$ has a uniform distribution over the interval $[0,1]$, and that therefore $-\log F(\mathbf{X}_i;\theta)$ has an exponential distribution $f(u) = e^{-u}$, $u > 0$.

(b) Show that the CDF of the entire sample, $\prod_{i=1}^n F(\mathbf{X}_i;\theta)$ is a pivot for θ. (Hint: the product of monotone functions is monotone).

(c) Show that a $(1-\alpha)\cdot 100\%$ confidence interval for θ can be constructed since $F(x;\theta)$ is monotone in θ using the result of part (b). (Hint: the sum of n i.i.d. exponential r.v.s with distribution $f(u) = e^{-u}$, $u > 0$ has a Gamma density).

10.4 Use the approach of the previous problem to construct $(1-\alpha)100\%$ confidence intervals for the following parameters.

(a) θ is the parameter in the density $f(x;\theta) = 2\theta x + 1 - \theta$, $0 \leq x \leq 1$, $-1 \leq \theta \leq 1$. Verify your results by numerical simulation for $n = 10$ using Matlab. Note it may be helpful to use Matlab's polynomial rooting procedure `roots.m` to find the interval endpoints. Note for this example you *cannot* use double sided GLRT since the GLRT is degenerate.

(b) θ is the median of the Cauchy distribution $f(x_i,\theta) = (1+(x-\theta)^2)/\pi$. Note that numerical integration may be required. Verify your results by numerical simulation for $n = 10$ using Matlab.

10.5 Let $\{x_1, \ldots, x_n\}$ be an i.i.d. sample of a Poisson r.v. with distribution $p_\theta(k) = P_\theta(x_i = k) = \frac{\theta^k}{k!}e^{-\theta}$, $k = 0, 1, 2, \ldots$. Use the GLRT derived in Exercise 8.4 to specify a $(1-\alpha)\%$ confidence interval on θ.

10.6 Let $\{X_i\}_{i=1}^n$ be i.i.d. following an exponential distribution

$$f(x;\theta) = \theta e^{-\theta x}, \quad x \geq 0$$

with $\theta > 0$.

(a) Derive the GLRT for the test of the hypotheses

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_1 &: \theta \neq \theta_0 \end{aligned}$$

with FA level α (Hint: the sum of n standard (mean = 1) exponential r.v.s is standard Gamma with parameter n).

(b) Using the results of (a) find a $(1-\alpha)\%$ confidence interval on θ.

End of chapter

11 SIGNAL DETECTION IN THE MULTIVARIATE GAUSSIAN MODEL

In this chapter we cover LR tests of simple hypotheses on the mean and covariance in the general multivariate Gaussian model. We will start with offline detection strategies when the measurement is a small dimensional vector of possibly correlated Gaussian observations. We then turn to online detection for change in mean and covariance of sampled Gaussian waveforms and this will bring the Kalman filter into the picture. This arises, for example, when we wish to decide on the mean or variance of Gaussian random process based on its time samples, a very common problem in signal processing, control and communications. While the focus is on simple hypotheses some discussion of unknown parameters is given.

Specifically, we will cover the following:

1. Offline methods:

* General vector Gaussian problem

* Detection of non-random signals in noise: matched-filter

* Detection of random signals in noise: filter-squarer and estimator-correlator

2. Online methods

* On line detection of non-random signals: causal matched-filter

* On-line detection for nonstationary signals: Kalman filter detector

11.1 OFFLINE METHODS

We have the following setup.

Observation: $\underline{X} = [X_1, \ldots, X_n]^T \sim \mathcal{N}_n(\underline{\mu}, \mathbf{R})$

mean: $\underline{\mu} = E[\underline{X}] = [\mu_1, \ldots, \mu_n]^T$

covariance: $\mathbf{R} = ((\mathrm{cov}(x_i, x_j)))_{i,j=1,\ldots,n}$

$$\mathbf{R} = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1,n} \\ \vdots & \ddots & \vdots \\ \sigma_{n,1} & \ldots & \sigma_n^2 \end{bmatrix}$$

Joint density

$$f(\underline{x};\ \underline{\mu}, \mathbf{R}) = \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{R}|}} \exp\left(-\tfrac{1}{2}(\underline{x}-\underline{\mu})^T \mathbf{R}^{-1} (\underline{x}-\underline{\mu})\right)$$

Consider simple detection problem

$$\begin{aligned} H_0 &: \underline{\mu} = \underline{\mu}_0, \quad \mathbf{R} = \mathbf{R}_0 \\ H_1 &: \underline{\mu} = \underline{\mu}_1, \quad \mathbf{R} = \mathbf{R}_1 \end{aligned}$$

Likelihood ratio is

$$\Lambda(\underline{x}) = \frac{\sqrt{|\mathbf{R}_0|}\exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu}_1)^T\mathbf{R}_1^{-1}(\underline{x}-\underline{\mu}_1)\right)}{\sqrt{|\mathbf{R}_1|}\exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu}_0)^T\mathbf{R}_0^{-1}(\underline{x}-\underline{\mu}_0)\right)}$$

Giving LRT

$$T(\underline{x}) =$$

$$\tfrac{1}{2}(\underline{x}-\underline{\mu}_0)^T\mathbf{R}_0^{-1}(\underline{x}-\underline{\mu}_0) - \tfrac{1}{2}(\underline{x}-\underline{\mu}_1)^T\mathbf{R}_1^{-1}(\underline{x}-\underline{\mu}_1) \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

where

$$\gamma = \log\eta + \tfrac{1}{2}\log\frac{|\mathbf{R}_1|}{|\mathbf{R}_0|}$$

Interpretation of LRT as distorted "Mahalanobis" distance test:

Define two norms in $\mathbb{R}^n$

$$\|\underline{z}\|_{R_0} = \underline{z}^T\mathbf{R}_0^{-1}\underline{z}, \qquad \|\underline{z}\|_{R_1} = \underline{z}^T\mathbf{R}_1^{-1}\underline{z},$$

Norm $\|\underline{z}\|_R$ emphasizes components of $\underline{z}$ which are colinear to eigenvectors of $\mathbf{R}$ associated with small eigenvalues

Then MP-LRT takes form of a comparison between weighted distances of $\underline{x}$ to $\underline{\mu}_0$ vs. $\underline{\mu}_1$

$$\|\underline{x}-\underline{\mu}_0\|^2_{R_0} - \|\underline{x}-\underline{\mu}_1\|^2_{R_1} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma'$$

An alternative form of LRT is the quadratic test

$$T'(\underline{x}) = \tfrac{1}{2}\underline{x}^T[\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}]\underline{x} + (\underline{\mu}_1^T\mathbf{R}_1^{-1} - \underline{\mu}_0^T\mathbf{R}_0^{-1})\underline{x} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma'$$

$$\gamma' = \log\eta + \tfrac{1}{2}\log\frac{|\mathbf{R}_1|}{|\mathbf{R}_0|} + \underline{\mu}_1^T\mathbf{R}_1^{-1}\underline{\mu}_1 - \underline{\mu}_0^T\mathbf{R}_0^{-1}\underline{\mu}_0$$

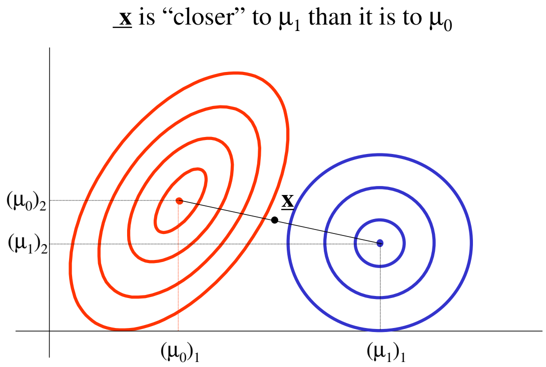


Figure 146: *LRT for general Gaussian problem compares "closeness" of $\underline{x}$ to distorted neighborhoods of the means μ_0 and μ_1.*

11.1.1 GENERAL CHARACTERIZATION OF LRT DECISION REGIONS

Divide treatement into four cases:

1. $\mathbf{R}_0 = \mathbf{R}_1$,
2. $\mathbf{R}_0 - \mathbf{R}_1 > 0$,
3. $\mathbf{R}_0 - \mathbf{R}_1 < 0$,
4. $\mathbf{R}_0 - \mathbf{R}_1$ non-singular

Case 1. $\mathbf{R}_0 = \mathbf{R}_1 = \mathbf{R}$:

In this case $T'(\underline{x}) = \underline{a}^T \underline{x}$ is linear function

$$\underline{a} = \mathbf{R}^{-1}(\underline{\mu}_1 - \underline{\mu}_0),$$

and decision regions are separated by a hyperplane.

Case 2. $\mathbf{R}_0 - \mathbf{R}_1 > 0$: (p.d.)

In this case, as $\mathbf{R}_0 > \mathbf{R}_1$ implies $\mathbf{R}_0^{-1} < \mathbf{R}_1^{-1}$,

$\mathbf{R}_1^{-1} - \mathbf{R}_0^{-1} = \Delta_{10}\mathbf{R}^{-1} > 0$ (p.d.):

and

$$T'(\underline{x}) = -\tfrac{1}{2}(\underline{x} - \underline{b})^T \Delta_{10}\mathbf{R}^{-1}(\underline{x} - \underline{b}) + c$$

$$\underline{b} = (\Delta_{10}\mathbf{R}^{-1})^{-1}[\mathbf{R}_1^{-1}\underline{\mu}_1 - \mathbf{R}_0^{-1}\underline{\mu}_0]$$

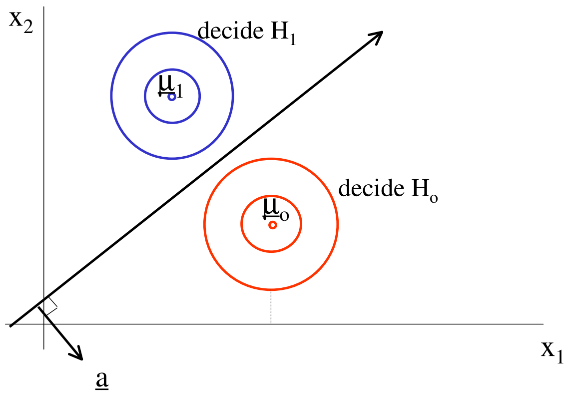


Figure 147: *For equal covariances of a multivariate Gaussian sample the decision regions are separated by a hyperplane (here shown for $\gamma^{'} = 0$ for which a is orthogonal to separating hyperplane).*

Hence the H_1 decision region is an ellipsoid

$$\mathcal{X}_1 = \{\tfrac{1}{2}(\underline{x} - \underline{b})^T [\Delta_{10}\mathbf{R}^{-1}](\underline{x} - \underline{b}) < \gamma^{''}\}$$

Case 3. $\mathbf{R}_0 < \mathbf{R}_1$

In this case

$\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1} = \Delta_{01}\mathbf{R}^{-1} > 0$ (p.d.):

and

$$T^{'}(\underline{x}) = \tfrac{1}{2}(\underline{x} - \underline{b})^T \ [\Delta_{01}\mathbf{R}^{-1}] \ (\underline{x} - \underline{b}) + c$$

$$\underline{b} \quad = \quad (\Delta_{01}\mathbf{R}^{-1})^{-1}[\mathbf{R}_0^{-1}\underline{\mu}_0 - \mathbf{R}_1^{-1}\underline{\mu}_1]$$

So now the H_0 decision region is an ellipsoid

$$\mathcal{X}_0 = \{\tfrac{1}{2}(\underline{x} - \underline{b})^T \Delta_{01}\mathbf{R}^{-1}(\underline{x} - \underline{b}) < \gamma^{''}\}$$

4. $\mathbf{R}_0 - \mathbf{R}_1$ not p.d, n.d., or singular

Let $\Delta_{01}\mathbf{R}^{-1}$ be defined as above

$$\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1} =: \Delta_{01}\mathbf{R}^{-1}$$

Let $\{\lambda_i\}_{i=1}^n$ denote the eigenvalues of this matrix

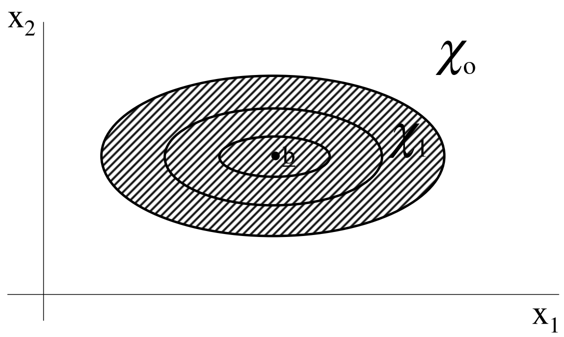


Figure 148: *For $\mathbf{R}_0 > \mathbf{R}_1$ the H_1 decision region is the interior of an ellipsoid for testing covariance of a multivariate Gaussian sample.*

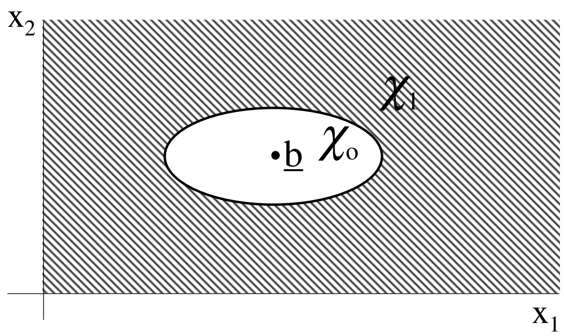


Figure 149: *For $\mathbf{R}_0 < \mathbf{R}_1$ the H_1 decision region for testing the covariance of a multivariate Gaussian sample is the exterior of an ellipsoid.*

Let the *signature* of $\Delta_{01}\mathbf{R}^{-1}$ be the binary vector

$$\underline{b} = [b_1, \ldots, b_n]^T$$

Definition: The signature of a non-singular symmetric matrix B are the signs of its eigenvalues arranged in decreasing order of magnitude.

$\Rightarrow$ If the signature $[b_1, \ldots, b_n]^T$ equals $[1, \ldots, 1]$ then all eigenvalues are positive and B is positive definite.

$\Rightarrow$ If $[b_1, \ldots, b_n]^T$ equals $[-1, \ldots, -1]$ then $-B$ is positive definite.

General representation:

$$[b_1, \ldots, b_n] = [\mathrm{sgn}(\lambda_1), \ldots, \mathrm{sgn}(\lambda_n)]$$

where

$$\mathrm{sgn}(u) := \begin{cases} 1, & u > 0 \\ 0, & u = 0 \\ -1 & u < 0 \end{cases}$$

We can rewrite the LRT test statistic as

$$\begin{aligned} T'(\underline{x}) &= \tfrac{1}{2}(\underline{x} - \underline{b})^T[\Delta_{01}\mathbf{R}^{-1}](\underline{x} - \underline{b}) + c \\ &= b_1\, z_1^2 + \ldots + b_n\, z_n^2 + c \end{aligned}$$

where

$$z_i = \lambda_i(\underline{x} - \underline{b})^T \underline{\nu}_i$$

and $\underline{\nu}_i$'s are eigenvectors of $\Delta_{01}\mathbf{R}^{-1}$.

Thus each decision region is hyperbolic.

11.1.2 CASE OF EQUAL COVARIANCES

Here $\mathbf{R}_0 = \mathbf{R}_1 = \mathbf{R}$ and LRT collapses to linear test

$$T(\underline{x}) = \Delta\underline{\mu}^T\mathbf{R}^{-1}\underline{x} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_1$$

where $\Delta\underline{\mu} = (\underline{\mu}_1 - \underline{\mu}_0)$

and

$\gamma_1 = \log\eta + \underline{\mu}_0^T\mathbf{R}^{-1}\underline{\mu}_0 - \underline{\mu}_1^T\mathbf{R}^{-1}\underline{\mu}_1$

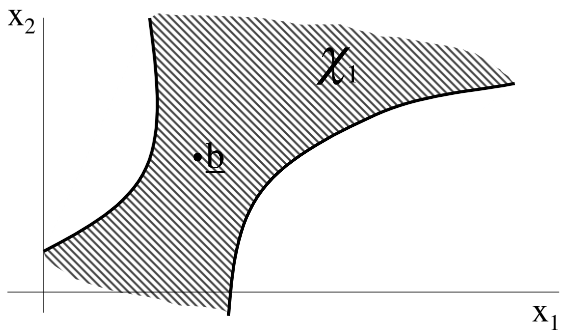


Figure 150: *For* $\mathbf{R}_0 - \mathbf{R}_1$ *non-singular but neither p.d. nor n.d. the* H_1 *decision region for testing the covariance of a multivariate Gaussian sample is a hyperboloid.*

DETECTOR PERFORMANCE

As the test statistic $T(\underline{X})$ is Gaussian suffices to find means

$$
\begin{aligned}
E_0[T] &= \Delta\underline{\mu}^T \mathbf{R}^{-1} \underline{\mu}_0 \\
E_1[T] &= \Delta\underline{\mu}^T \mathbf{R}^{-1} \underline{\mu}_1
\end{aligned}
$$

and variances

$$
\begin{aligned}
\mathrm{var}_0(T) &= \mathrm{var}_0\left(\Delta\underline{\mu}^T \mathbf{R}^{-1} \underline{X}\right) \\
&= \Delta\underline{\mu}^T \mathbf{R}^{-1} \underbrace{\mathrm{cov}_0(\underline{X})}_{R} \mathbf{R}^{-1} \Delta\underline{\mu} \\
&= \Delta\underline{\mu}^T \mathbf{R}^{-1} \Delta\underline{\mu} \\
\mathrm{var}_1(T) &= \mathrm{var}_0(T)
\end{aligned}
$$

Thus we find

$$
P_F = \alpha = 1 - \mathcal{N}\left(\frac{\gamma_1 - E_0[T]}{\sqrt{\mathrm{var}_0(T)}}\right)
$$

so that the NP MP-LRT test is

$$\Delta\underline{\mu}^T \mathbf{R}^{-1} \underline{x} \overset{H_1}{\underset{H_0}{\gtrless}} \sqrt{\Delta\underline{\mu}^T \mathbf{R}^{-1} \Delta\underline{\mu}}\, \mathcal{N}^{-1}(1-\alpha) \ + \ \Delta\underline{\mu}^T \mathbf{R}^{-1} \underline{\mu}_0$$

or equivalently

$$\frac{\Delta\underline{\mu}^T \mathbf{R}^{-1}(\underline{x} - \underline{\mu}_0)}{\sqrt{\Delta\underline{\mu}^T \mathbf{R}^{-1} \Delta\underline{\mu}}} \overset{H_1}{\underset{H_0}{\gtrless}} \mathcal{N}^{-1}(1-\alpha)$$

NOTES:

1. For $\underline{\mu}_0 \neq 0$: MP test is not UMP w.r.t unknown parameter variations
2. For $\underline{\mu}_0 = 0$: MP test is UMP w.r.t. constant positive scaling of $\underline{\mu}_1$

Next find power:

$$P_D = \beta = 1 - \mathcal{N}\left(\frac{\gamma_1 - E_1[T]}{\sqrt{\mathrm{var}_1(T)}}\right)$$

giving ROC curve:

$$\beta = 1 - \mathcal{N}\left(\mathcal{N}^{-1}(1-\alpha) - d\right)$$

where d is detectability index

$$\begin{aligned} d &= \frac{E_1[T] - E_0[T]}{\sqrt{\mathrm{var}_0(T)}} \\ &= \sqrt{\Delta\underline{\mu}^T \mathbf{R}^{-1} \Delta\underline{\mu}} \end{aligned}$$

Example 44 *Detection of known signal in white noise*

$$\begin{aligned} H_0 : x_k &= w_k \\ & & k = 1, \ldots, n \\ H_1 : x_k &= s_k + w_k \end{aligned}$$

* $\underline{w} \sim \mathcal{N}_n(0, \sigma^2 \mathbf{I})$,

* $\underline{s}$ and σ^2 known

Identify:

$$\underline{\mu}_0 = 0, \quad \Delta\underline{\mu} = \underline{s}, \qquad \mathbf{R} = \sigma^2 \mathbf{I}$$

so that the LRT takes the form of a matched filter

$$T(x) = \underline{s}^T \underline{x} = \sum_{k=1}^{n} s_k x_k \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

$\gamma = \sigma^2 \log \eta - \|\underline{s}\|^2$

GEOMETRIC INTERPRETATION

The LRT can be expressed geometrically as a signal "projection" detector

Projection of $\underline{x}$ onto $\underline{s}$ is

$$\begin{aligned} \hat{\underline{x}} &= \underbrace{\left[\frac{\underline{s}\, \underline{s}^T}{\|\underline{s}\|^2}\right]}_{\Pi_s} \underline{x} \\ &= \underline{s}\, \frac{\underline{s}^T \underline{x}}{\|\underline{s}\|^2} \\ &= \underline{s}\, \underbrace{\frac{< \underline{s}, \underline{x} >}{\|\underline{s}\|^2}}_{\text{Proj. coef.}} \end{aligned}$$

Length of this projection is

$$\begin{aligned} \|\hat{\underline{x}}\| &= \|\underline{s}\| \left| \frac{\underline{s}^T \underline{x}}{\|\underline{s}\|^2} \right| \\ &= |T(\underline{x})| \, \frac{1}{\|\underline{s}\|} \end{aligned}$$

Conclude:

* LRT is a threshold test on the projection coefficient of the orthogonal projection of $\underline{x}$ onto $\underline{s}$

* LRT is threshold test on "signed length" of $\hat{\underline{x}}$

* LRT is related to LLS estimator $\hat{\underline{x}}$ of $\underline{x}$ given $\underline{s}$

PERFORMANCE

Equivalently, for MP test of level α we have

$$\frac{\underline{s}^T \, \underline{x}}{\|\underline{s}\| \, \sigma} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \mathcal{N}^{-1}(1-\alpha)$$

This test is UMP relative to signal energy $\|\underline{s}\|^2$

Now compute detectability index:

$$d^2 = \|\underline{s}\|^2/\sigma^2 = \frac{\sum_{k=1}^n s_k^2}{\sigma^2} =: \text{SNR} \qquad (128)$$

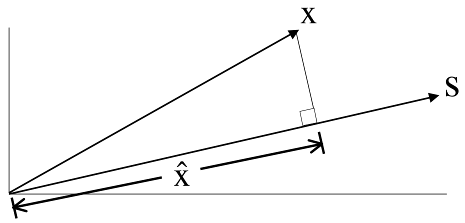


Figure 151: *MP detector applies a threshold to projection coefficient* $<\underline{x}, \underline{s}>/\|\underline{s}\|^2$ *of orthogonal projection of* $\underline{x}$ *onto* $\underline{s}$*, shown here for the case of* $n = 2$.

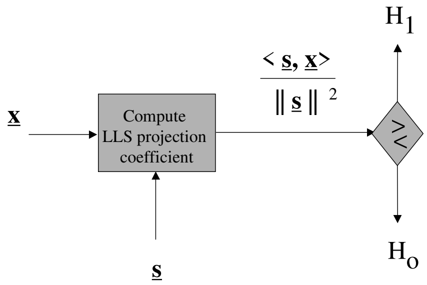


Figure 152: *MP detector block diagram implemented with a LLS estimator of* $\underline{x}$.

NOTE:

* detection index is invariant to shape of waveform $\underline{s}$.

* detector power only depends on total signal-to-noise ratio (SNR).

Note: two ways to implement optimal detector

Cross-correlator:

Matched Filter:

Example 45 *Detection of known signal in non-white noise:*

$$\begin{aligned} H_0 : x_k &= w_k \\ & \qquad\qquad k = 1, \ldots, n \\ H_1 : x_k &= s_k + w_k \end{aligned}$$

* $\underline{w} \sim \mathcal{N}_n(0, \mathbf{R})$, $\mathbf{R}$ not scaled identity.

Optimal detector

$$T(\underline{x}) = \frac{\underline{s}^T \mathbf{R}^{-1} \underline{x}}{\sqrt{\underline{s}^T \mathbf{R}^{-1} \underline{s}}} \underset{H_0}{\overset{H_1}{\gtrless}} \mathcal{N}^{-1}(1-\alpha)$$

Q. How to modularize detector?

A. transform to the white noise case via preprocessing with matrix $\mathbf{H}$

Produces white noise measurements

$$\underline{\tilde{x}} = \mathbf{H} \cdot \underline{x}$$

We will require matrix filter H to have properties:

1. $\text{cov}_0(\underline{\tilde{x}}) = \text{cov}_1(\underline{\tilde{x}}) = \mathbf{I}$: $\Rightarrow$ whitening property

2. H is invertible matrix: $\Rightarrow$ output remains sufficient statistic

MATRIX FACTORIZATION

For any symmetric positive definite covariance matrix $\mathbf{R}$ there exists a positive definite square root factor $\mathbf{R}^{\frac{1}{2}}$ and a positive definite inverse factor $\mathbf{R}^{-\frac{1}{2}}$ which satisfy:

$$\mathbf{R} = \mathbf{R}^{\frac{1}{2}} \mathbf{R}^{\frac{1}{2}}, \text{ and } \mathbf{R}^{-1} = \mathbf{R}^{-\frac{1}{2}} \mathbf{R}^{-\frac{1}{2}}.$$

There are many possible factorizations of this type. We have already seen the Cholesky factorization in Chapter 6 which yields upper and lower triangular factors. Here we focus on a symmetric factorization given by the eigendecomposition of $\mathbf{R} = \mathbf{U}\mathbf{D}\mathbf{U}^T$, where

* $\mathbf{D} = \text{diag}(\lambda_i)$ are (positive) eigenvalues of $\mathbf{R}$

* $\mathbf{U} = [\underline{\nu}_1, \ldots, \underline{\nu}_p]$ are (orthogonal) eigenvectors of $\mathbf{R}$

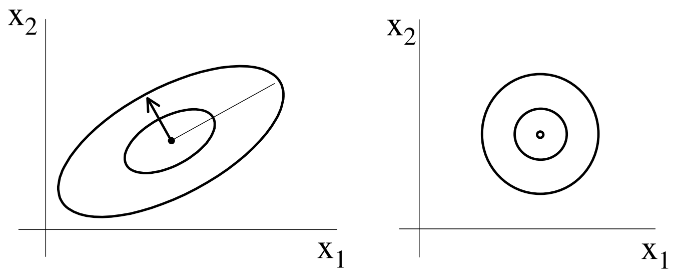


Figure 153: *Matrix prewhitener H applied to $\underline{x}$ renders the transforms contours of the multivariate Gaussian density to concentric circles (spherically symmetric).*

As $\mathbf{U}^T\mathbf{U} = \mathbf{I}$

$$\mathbf{R} = \mathbf{U}\mathbf{D}\mathbf{U}^T = \mathbf{U}\mathbf{D}^{\frac{1}{2}}\mathbf{D}^{\frac{1}{2}}\mathbf{U}^T = \mathbf{U}\mathbf{D}^{\frac{1}{2}}\mathbf{U}^T\ \mathbf{U}\mathbf{D}^{\frac{1}{2}}\mathbf{U}^T$$

Therefore we can identify

$$\mathbf{R}^{\frac{1}{2}} = \mathbf{U}\mathbf{D}^{\frac{1}{2}}\mathbf{U}^T.$$

Furthermore, since $\mathbf{U}^{-1} = \mathbf{U}^T$ we have

$$\mathbf{R}^{-\frac{1}{2}} = \mathbf{U}\mathbf{D}^{-\frac{1}{2}}\mathbf{U}^T$$

which satisfy the desired properties of square root factors and are in addition symmetric.

Using square root factors the test statistic can be rewritten as

$$\begin{aligned} T(\underline{x}) &= \frac{\underline{s}^T\mathbf{R}^{-\frac{1}{2}}\mathbf{R}^{-\frac{1}{2}}\underline{x}}{\sqrt{\underline{s}^T\mathbf{R}^{-\frac{1}{2}}\mathbf{R}^{-\frac{1}{2}}\underline{s}}} \\ &= \frac{\underline{\tilde{s}}^T\underline{\tilde{x}}}{\|\underline{\tilde{s}}\|} \end{aligned}$$

Where $\underline{x}$, $\underline{s}$ are transformed vectors

$$\underline{\tilde{x}} = \mathbf{R}^{-\frac{1}{2}}\underline{x}, \quad \underline{\tilde{s}} = \mathbf{R}^{-\frac{1}{2}}\underline{s}$$

Now we see that

$E_0[\underline{\tilde{X}}] = 0, \quad E_1[\underline{\tilde{X}}] = \underline{\tilde{s}}, \quad \text{cov}_0(\underline{\tilde{X}}) = \text{cov}_1(\underline{\tilde{X}}) = \mathbf{I}$

so that problem is equivalent to testing for a signal in white noise of unit variance.

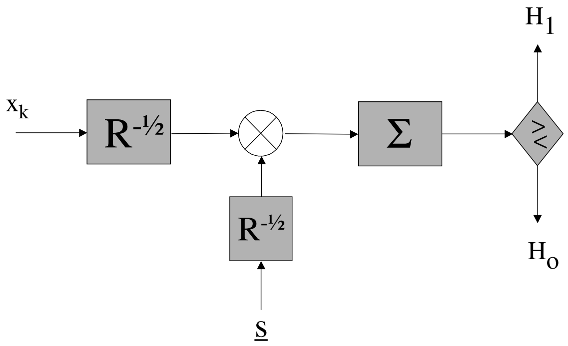


Figure 154: *Matrix prewhitener $H = \mathbf{R}^{-\frac{1}{2}}$ is applied prior to optimal matched filter detection.*

Detectability index for non-white noise:

Note: $d^2 = \|\underline{\tilde{s}}\|^2 = \underline{s}^T \mathbf{R}^{-1} \underline{s}$.

Remark

No longer is detection performance independent of shape of $\underline{s}$

OPTIMAL SIGNAL DESIGN FOR NON-WHITE NOISE:

Constraint: $\|\underline{s}\|^2 = 1$

Maximize: $d^2 = \underline{s}^T \mathbf{R}^{-1} \underline{s}$

Solution: Rayleigh quotient theorem specifies:

$$\frac{\underline{s}^T \mathbf{R}^{-1} \underline{s}}{\underline{s}^T \underline{s}} \leq \frac{1}{\min_i \lambda_i^R}$$

λ_i^R = an eigenvalue of $\mathbf{R}$.

Furthermore

$$\frac{\underline{s}^T \mathbf{R}^{-1} \underline{s}}{\underline{s}^T \underline{s}} = \frac{1}{\min_i \lambda_i^R}$$

when $\underline{s}$ is (any) minimizing eigenvector of $\mathbf{R}$ (there will be multiple minimizing eigenvectors if more than one eigenvalue $\{\lambda_k\}$ equals $\min_i \lambda_i^R$). The intuition here is that the best signal vector points in the direction of signal space that has the lowest noise power; hence maximizing the SNR over the set of fixed energy signals.

Example 46 *Application: (Real) Signal detection in a sensor array*

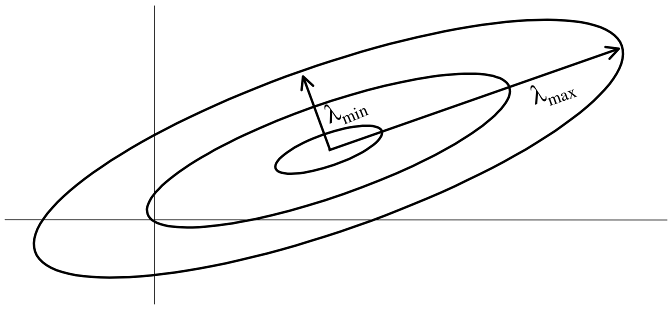


Figure 155: *The optimal signal which maximizes detectability is the eigenvector of noise covariance* $\mathbf{R}$ *with minimum eigenvalue.*

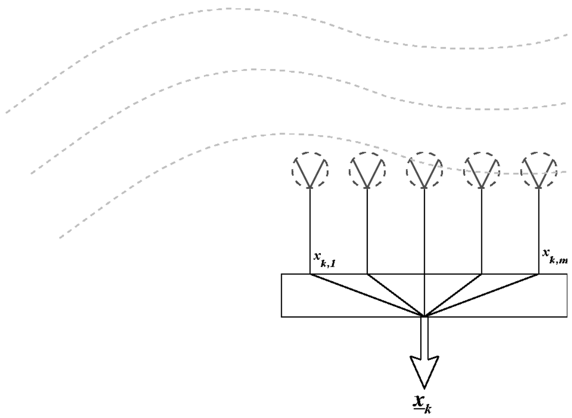


Figure 156: *Sensor array receiving signal wavefront generates spatio-temporal measurement.*

k-th snapshot of p-sensor array output is a *multi-channel* measurement:

$$\underline{x}_k = \underline{a}\, s + \underline{v}_k, \quad k = 1, \ldots, n$$

or equivalently we have $p \times n$ measurement matrix

$$\mathbf{X} = [\underline{x}_1, \ldots, \underline{x}_n]$$

* $\underline{a}$: array response vector
* $\underline{v}_k$: Gaussian $\mathcal{N}_p(0, \mathbf{R})$, known spatial covariance $\mathbf{R}$
* s: deterministic signal amplitude

Three cases of interest:

1. Detection of known signal amplitude
2. Detection of positive signal amplitude
3. Detection of non-zero signal amplitude

Case 1: Known signal amplitude

$$\begin{aligned} H_0 : s &= 0, \quad k = 1, \ldots, n \\ H_1 : s &= s_1, \quad k = 1, \ldots, n \end{aligned}$$

Approach: reduce to single-channel problem via coordinate rotation

As $\underline{a}$, $\mathbf{R}$ are known, we can transform the array to one with

* spatially uncorrelated noise ($\mathbf{R}$ diagonal)
* signal energy present only in first channel.

Define the $p \times p$ matrix $\mathbf{H}$:

$$\mathbf{H} = \left[\frac{1}{\tilde{a}}\, \mathbf{R}^{-\frac{1}{2}} \underline{a}, \underline{\nu}_2, \ldots, \underline{\nu}_p\right]$$

where

* $\tilde{a} = \sqrt{\underline{a}^T \mathbf{R}^{-1} \underline{a}}$
* $\underline{\nu}_i$ orthonormal vectors orthogonal to $\mathbf{R}^{-\frac{1}{2}}\underline{a}$ (found via Gramm-Schmidt)

Then

$$\underbrace{\mathbf{H}^T \mathbf{R}^{-\frac{1}{2}}}_{\mathbf{W}}\, \underline{a} = \begin{bmatrix} \tilde{a} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \tilde{a}\, \underline{e}_1$$

Now as $\mathbf{W} = \mathbf{H}^T \mathbf{R}^{-\frac{1}{2}}$ is invertible, the following is equivalent measurement

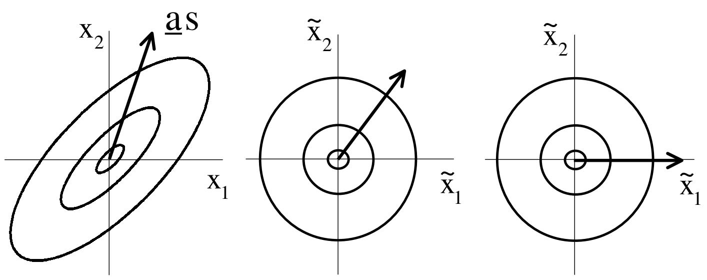


Figure 157: *Transformation of Gaussian multichannel problem to a Gaussian single channel problem is a two step procedure. First a whitening coordinate transformation* $\mathbf{R}^{-\frac{1}{2}}$ *is applied to measurements* $\underline{x} = \underline{a}s + \underline{n}$ *(joint density in original coordinates is shown in left panel) which makes noise component* $\underline{n}$ *i.i.d. (transformed measurements have joint density with spherically symmetric constant contours shown in middle panel). Then a pure rotation (unitary matrix) H is applied to the transformed measurements* $\underline{\tilde{x}}$ *which aligns its signal component* $\mathbf{R}^{-\frac{1}{2}}\underline{a}s$ *with the first coordinate axis (right panel).*

$$
\begin{aligned}
\underline{\tilde{X}}_k &= \mathbf{W}\underline{X}_k, \\
&= s\mathbf{W}\underline{a} + \mathbf{W}\underline{V}_k, \\
&= s_1 \tilde{a}\ \underline{e}_1 + \underline{\tilde{V}}_k
\end{aligned}
$$

where $\underline{\tilde{V}}_k$'s are i.i.d. zero mean Gaussian with identity covariance

$$\mathrm{cov}(\underline{\tilde{V}}_k) = \mathbf{W}\mathbf{R}\mathbf{W}^T = \mathbf{H}^T\ \mathbf{R}^{-\frac{1}{2}}\mathbf{R}\mathbf{R}^{-\frac{1}{2}}\ \mathbf{H} = \mathbf{H}^T\mathbf{H} = \mathbf{I}$$

Matrix representation

$$\tilde{\mathbf{X}} = s_1\ \tilde{a}\underline{1}^T + \mathbf{V}$$

where $\mathbf{V}$ is a $p \times n$ matrix of i.i.d. $\mathcal{N}(0,1)$'s

Note properties:

* all rows of $\tilde{\mathbf{X}} = \mathbf{W}\mathbf{X}$ are independent

* only first row $\tilde{\mathbf{X}}_{1*} = \underline{e}_1^T\tilde{\mathbf{X}}$ depends on s_1

Therefore LR only depends on the first row $\tilde{\mathbf{X}}$

$$
\begin{aligned}
\Lambda(\tilde{\mathbf{X}}) &= \frac{f(\tilde{\mathbf{X}}_{1*}, \tilde{\mathbf{X}}_{2*}, \ldots, \tilde{\mathbf{X}}_{n*};\ s = s_1)}{f(\tilde{\mathbf{X}}_{1*}, \tilde{\mathbf{X}}_{2*}, \ldots, \tilde{\mathbf{X}}_{n*};\ s = 0)} \\
&= \frac{f(\tilde{\mathbf{X}}_{1*};\ s = s_1)}{f(\tilde{\mathbf{X}}_{1*};\ s = 0)} \underbrace{\prod_{i=2}^{p} \frac{f(\tilde{\mathbf{X}}_{i*};\ s = s_1)}{f(\tilde{\mathbf{X}}_{i*};\ s = 0)}}_{=1} \\
&= \frac{f(\tilde{\mathbf{X}}_{1*};\ s = s_1)}{f(\tilde{\mathbf{X}}_{1*};\ s = 0)} = \Lambda(\tilde{\mathbf{X}}_{1*})
\end{aligned}
$$

Thus we have reduced the problem to equivalent hypotheses that a (row) vector measurement $\underline{z}^T = \tilde{\mathbf{X}}_{1*}$ contains a constant signal in i.i.d. Gaussian noise of variance 1

$$
\begin{array}{lcl}
H_0 : \underline{z} = \underline{\tilde{v}} & & H_0 : z_k = \tilde{v}_k \\
& \Leftrightarrow & \\
H_1 : \underline{z} = s_1\ \tilde{a}\ \underline{1} + \underline{\tilde{v}} & & H_1 : z_k = s_1\ \tilde{a} + \tilde{v}_k
\end{array}
$$

The LRT follows immediately from our previous work in detection of constant signal $\mu = s_1 \tilde{a}$

$$
s_1 \tilde{a}\ \overline{z_i} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma
$$

Or, as $\tilde{a}$ is positive, the final form of the LRT is

$$
\begin{aligned}
T(z) = \sqrt{n}\ \overline{z_i} \quad & \overset{H_1}{\underset{H_0}{\gtrless}} \quad \mathcal{N}^{-1}(1-\alpha) \quad s_1 > 0, \\
\sqrt{n}\ \overline{z_i} \quad & \overset{H_0}{\underset{H_1}{\gtrless}} \quad -\mathcal{N}^{-1}(1-\alpha) \quad s_1 < 0,
\end{aligned}
$$

The power of the test is determined by the detectibility index

$$
d = \frac{|E[\overline{Z_i}|H_1]|}{\sqrt{\mathrm{var}(\overline{Z_i}|H_0)}} = \sqrt{n}\ |s_1 \tilde{a}| = \sqrt{n}\ |s_1|\ \sqrt{\underline{a}^T \mathbf{R}^{-1} \underline{a}}
$$

We can express LRT in original coordinates by identifying

$$
\begin{aligned}
\underline{z}^T &= \tilde{\mathbf{X}}_{1*} = \underline{e}_1^T \tilde{\mathbf{X}} = \underline{e}_1^T \underbrace{\mathbf{W}}_{\mathbf{H}^T \mathbf{R}^{-\frac{1}{2}}} \mathbf{X} \\
&= \underbrace{\frac{1}{\sqrt{\underline{a}^T \mathbf{R}^{-1} \underline{a}}}\, \underline{a}^T \mathbf{R}^{-\frac{1}{2}}}_{\underline{e}_1^T \mathbf{H}^T} \mathbf{R}^{-\frac{1}{2}} \mathbf{X} \\
&= \frac{1}{\sqrt{\underline{a}^T \mathbf{R}^{-1} \underline{a}}}\, \underline{a}^T \mathbf{R}^{-1} \mathbf{X}
\end{aligned}
$$

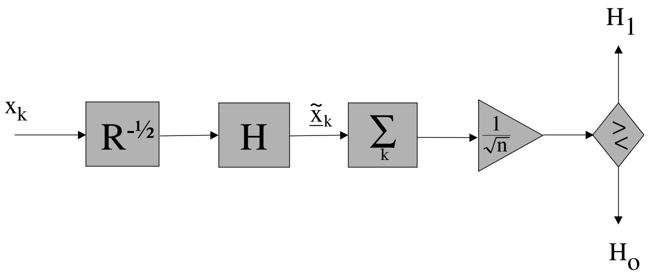


Figure 158: *LRT for detecting presence of a spatio-temporal signal implemented with whitening and coordinate rotation preprocessing.*

and the identity

$$\overline{z_i} = (\underline{z}^T \underline{1}) \frac{1}{n}$$

to obtain ($s_1 > 0$)

$$T(\underline{z}) = \frac{1}{\sqrt{n\underline{a}^T \mathbf{R}^{-1}\underline{a}}} \underline{a}^T \mathbf{R}^{-1} \mathbf{X} \underline{1} \overset{H_1}{\underset{H_0}{\gtrless}} \mathcal{N}^{-1}(1-\alpha),$$

OBSERVATIONS

1. The LRT above is UMP w.r.t. any positive amplitude s_1
2. A modified LRT is UMP w.r.t. any negative amplitude s_1
3. The detectibility index

$$d = \sqrt{n}|s_1| \underbrace{\sqrt{\underline{a}^T \mathbf{R}^{-1} \underline{a}}}_{\text{ASNR}}$$

depends on normalized array SNR = ASNR

$\Rightarrow$ ASNR depends only on $\|\underline{a}\|$ when noise $\underline{v}_k$ is spatially white ($\mathbf{R} = \sigma^2 \mathbf{I}$).

4. Coherent interferers can severely degrade performance

Case 2: Unknown signal amplitude

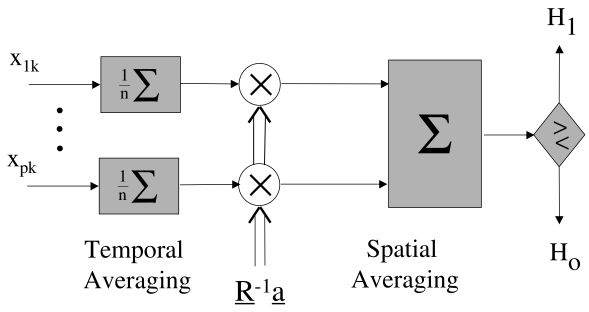


Figure 159: *LRT for detecting presence of a spatio-temporal signal implemented without coordinate transformation preprocessing.*

$$H_0 : s = 0, \quad k = 1, \ldots, n$$

$$H_1 : s \neq 0, \quad k = 1, \ldots, n$$

No UMP exists!

Solution: double sided GLRT

$$|T(z)| = \sqrt{n}|\overline{z_i}| = \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \mathcal{N}^{-1}(1-\alpha/2)$$

1. Implementation of GLRT via signal subspace projection:

Projection of $\underline{z}$ onto $\underline{s} = n^{-1}\, \underline{1}$ is

$$\begin{aligned} \hat{\underline{z}} &= \underbrace{\left[\frac{\underline{s}\,\underline{s}^T}{\|\underline{s}\|^2}\right]}_{\Pi_s} \underline{z} \\ &= \underline{s}\, \frac{\overbrace{\underline{s}^T \underline{z}}^{\overline{z_i}}}{\|\underline{s}\|^2} \end{aligned}$$

* Π_s = signal subspace projection operator

Length of $\hat{\underline{z}}$ is

$$\begin{aligned}
\|\hat{\underline{z}}\| &= \|\underline{s}\| \left| \frac{\underline{s}^T \underline{z}}{\|\underline{s}\|^2} \right| \\
&= |\overline{z_i}| \frac{1}{\|\underline{s}\|} \\
&= |\overline{z_i}|
\end{aligned}$$

Conclude:

* GLRT is a threshold test on the length of the orthogonal projection of $\underline{z}$ onto span($\underline{s}$)

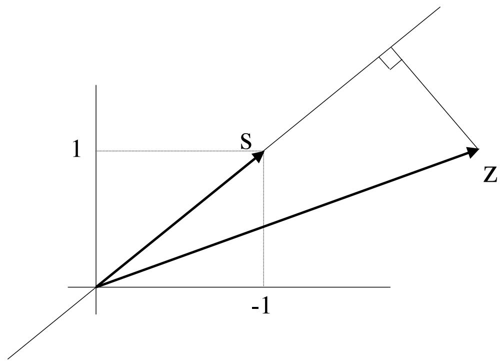


Figure 160: *GLRT detector thresholds length of orthogonal projection of $\underline{z}$ onto $\underline{s}$, shown here for the case of $n = 2$.*

2. Implementation of GLRT via "noise subspace" projection:

Recall orthogonal decomposition

$$\underline{z} = \Pi_s z + [\mathbf{I} - \Pi_s]\underline{z}$$

* Π_s = signal subspace projection operator

* $\mathbf{I} - \Pi_s$ = noise subspace projection operator

With this we can express GLRT as

$$|\overline{z_i}|^2 = \|\Pi_s \underline{z}\| = \|\underline{z}\|^2 - \| \underbrace{[\mathbf{I} - \Pi_s]\underline{z}}_{\underline{z}-\hat{\underline{z}}} \|^2 \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma'$$

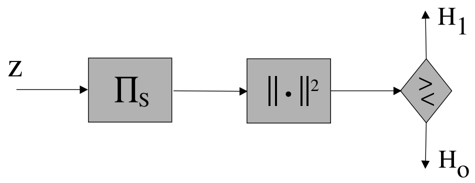


Figure 161: *GLRT detector block diagram implemented via signal subspace projection.*

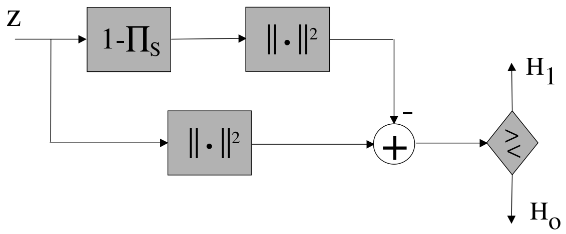


Figure 162: *GLRT detector block diagram implemented via noise subspace projection.*

11.1.3 CASE OF EQUAL MEANS, UNEQUAL COVARIANCES

Here $\underline{\mu}_0 = \underline{\mu}_1 = \underline{\mu}$ and LRT collapses to purely quadratic test

$$T(\underline{x}) = (\underline{x} - \underline{\mu})^T [\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}](\underline{x} - \underline{\mu}) \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

where $\gamma = 2\ln\eta$. Note that for convenience we have chosen to absorb the factor 1/2 in the log likelihood ratio into the threshold γ.

Analysis will be simplified by prefiltering to diagonalize $\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}$

$\Rightarrow$ Require prefilter to perform simultaneous diagonalization

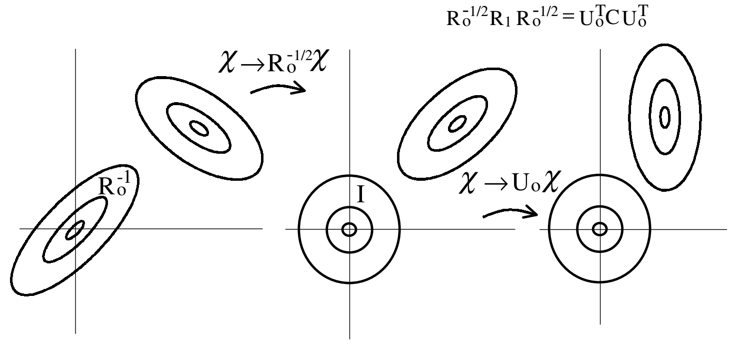


Figure 163: *Illustration of simultaneous whitening of two Gaussian data sets, or equivalently simultaneous diagonalization of two p.d. matrices as a two stage procedure. First one of the matrices is diagonalized and scaled to have identical diagonal elements via appropriate coordinate transformation (superposition of the constant contours of the two densities is shown on left panel along with the result of the coordinate transformation in middle panel). Then a unitary transformation is applied to diagonalize the other matrix without affecting the transformed first matrix (constant contours of the two densities shown on right panel).*

PERFORMANCE ANALYSIS

Under H_0 reexpress $(\underline{x} - \underline{\mu})^T [\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}](\underline{x} - \underline{\mu})$ as

$$T(\underline{X}) \;=\; \underbrace{(\underline{X} - \underline{\mu})^T \mathbf{R}_0^{-\frac{1}{2}}}_{=\underline{Z}^T \sim \mathcal{N}_n(0,\mathbf{I})} [\mathbf{I} - \mathbf{R}_0^{\frac{1}{2}} \mathbf{R}_1^{-1} \mathbf{R}_0^{\frac{1}{2}}] \underbrace{\mathbf{R}_0^{-\frac{1}{2}} (\underline{X} - \underline{\mu})}_{=\underline{Z} \sim \mathcal{N}_n(0,\mathbf{I})}$$

Now let $\mathbf{R}_0^{\frac{1}{2}} \mathbf{R}_1^{-1} \mathbf{R}_0^{\frac{1}{2}}$ have eigendecomposition

$$\mathbf{R}_0^{\frac{1}{2}} \mathbf{R}_1^{-1} \mathbf{R}_0^{\frac{1}{2}} = \mathbf{U}_0^T \mathbf{C} \mathbf{U}_0$$

* $\mathbf{U}_0$ orthogonal matrix of eigenvectors

* $\mathbf{C} = \text{diag}(c_1, \ldots, c_n)$ diagonal matrix of eigenvalues.

$$\begin{aligned} T(\underline{X}) &= \underbrace{(\mathbf{U}_0\underline{Z})^T}_{\mathcal{N}_n(0,\mathbf{I})}[\mathbf{I} - \mathbf{C}]\underbrace{(\mathbf{U}_0\underline{Z})}_{\mathcal{N}_n(0,\mathbf{I})} \\ &= n\overline{(1-c)}\underbrace{\sum_i \frac{Z_i^2(1-c_i)}{\sum_j(1-c_j)}}_{\text{Chi-sq.-mixture}} \end{aligned}$$

where

$$\overline{(1-c)} = n^{-1}\sum_i (1-c_i)$$

There are two cases to consider: $0 < c_i < 1$ vs $c_i > 1$. Note that consideration of $c_i = 0$ is not required since we have assumed that $\mathbf{R}_1$ and $\mathbf{R}_0$ are positive definite matrices.

CASE 1: $0 < c_i < 1$ for all i

Here

$$\overline{(1-c)} > 0$$

so we can absorb it into into threshold γ

This gives MP level α test in terms of orthogonalized measurements z_i

$$\sum_i \frac{z_i^2(1-c_i)}{\sum_j(1-c_j)} \underset{H_0}{\overset{H_1}{\gtrless}} \overline{\chi}_{n,1-c}^{-1}(1-\alpha)$$

Finally, retracing our steps to the original observables we have the implementable level α LRT test

$$(\underline{x} - \underline{\mu})^T(\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1})(\underline{x} - \underline{\mu}) \underset{H_0}{\overset{H_1}{\gtrless}} a\overline{\chi}_{n,1-c}^{-1}(1-\alpha).$$

Here $a = \sum_{i=1}^n (1-c_i)$ and $\overline{\chi}_{n,1-c}$ is the CDF of Chi-square-mixture r.v. with n degrees of freedom and mixture parameter vector

$$1 - c = [1 - c_1, \ldots, 1 - c_n]^T$$

(Johnson, Kotz and Balakrishnan [30, Sec. 18.8]).

It remains to find the power:

In a similar manner, under H_1 we can express

$$\begin{aligned} T(\underline{X}) &= (\underline{X} - \underline{\mu})^T\mathbf{R}_1^{-\frac{1}{2}}[\mathbf{R}_1^{\frac{1}{2}}\mathbf{R}_0^{-1}\mathbf{R}_1^{\frac{1}{2}} - \mathbf{I}]\mathbf{R}_1^{-\frac{1}{2}}(\underline{X} - \underline{\mu}) \\ &= (\mathbf{U}_1\underline{Z})^T[\mathbf{C}^{-1} - \mathbf{I}](\mathbf{U}_1\underline{Z}) \end{aligned}$$

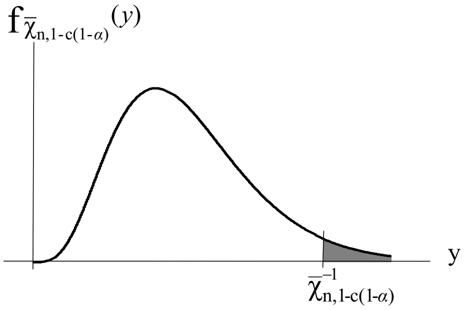


Figure 164: *For $c \leq c_i < 1$, threshold of test between two multivariate Gaussian models with identical means but unequal covariances is determined by quantile of Chi-square-mixture p.d.f.*

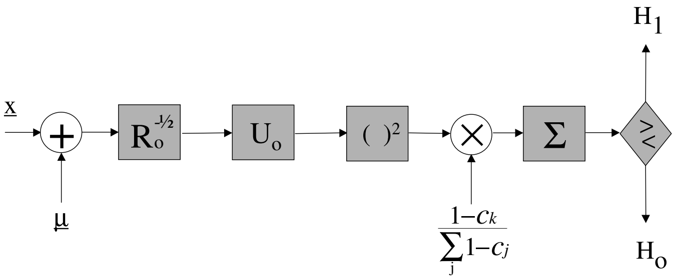


Figure 165: *An implementation of the MP-LRT for equal means unequal covariances using orthogonal prefilter $\mathbf{U}_0$ obtained from eigendecomposition: $\mathbf{R}_0^{\frac{1}{2}}\mathbf{R}_1^{-1}\mathbf{R}_0^{\frac{1}{2}} = \mathbf{U}_0^T\mathbf{C}\mathbf{U}_0$, where $\mathbf{C}$ is diagonal.*

$$= \overline{n(1/c-1)} \underbrace{\sum_i \frac{Z_i^2(1/c_i - 1)}{\sum_j (1/c_j - 1)}}_{Chi-sq.-mixture}$$

where $\mathbf{U}_1$ in the above is an orthogonal matrix.

As $\overline{(1/c-1)} > 0$, we easily obtain power as:

$$\beta = 1 - \overline{\chi}_{n,1/c-1}(\rho\, \overline{\chi}^{-1}_{n,1-c}(1-\alpha))$$

where

$$\rho = \overline{(1-c)}/\overline{(1/c-1)}$$

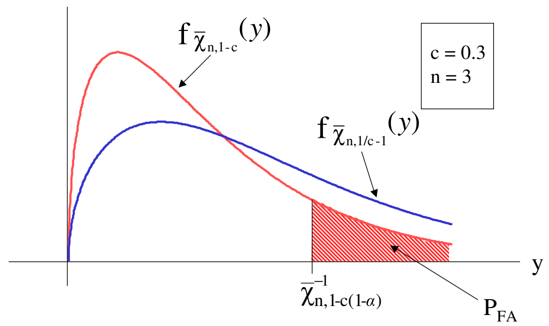


Figure 166: *For $c \leq c_i < 1$ ROC of test between two multivariate Gaussian models with identical means but unequal covariances is determined by upper quantiles of pair of Chi-square-mixture p.d.f.'s*

CASE 2: $c_i > 1$ for all i

Here we have $\overline{(1-c)} < 0$ and the constant in the test can be absorbed into threshold only with change of the inequalities.

Obtain the MP level α test in z_i coordinates

$$\sum_i \frac{z_i^2(1-c_i)}{\sum_j (1-c_j)} \overset{H_0}{\underset{H_1}{\gtrless}} \overline{\chi}^{-1}_{n,1-c}(\alpha)$$

and, using similar arguments as before, we obtain power curve

$$\beta = \overline{\chi}_{n,1/c-1}(\rho\, \overline{\chi}^{-1}_{n,1-c}(\alpha))$$

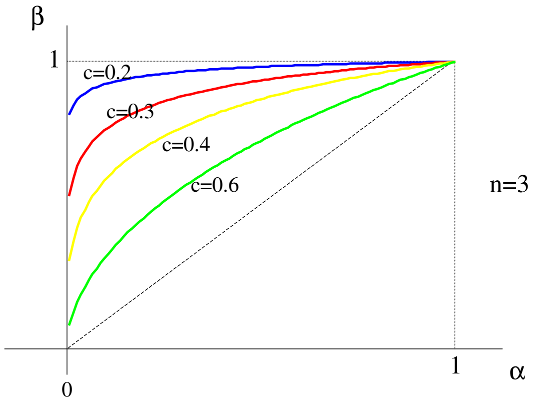


Figure 167: *ROC curve corresponding to Fig. 166 with various parameters c_1, c_2, c_3 for $n = 3$.*

Case 3, where some c_i's satsify the condition in Case 1 and others satisfy that of case 2 is more complicated as we end up with a Chi-squared difference in our test statistic.

11.2 APPLICATION: DETECTION OF RANDOM SIGNALS

Example 47 *Detection of shift in variance of white noise*

$$\begin{aligned} H_0 : x_k &= w_k \\ H_1 : x_k &= s_k + w_k \end{aligned}$$

$w_k \sim \mathcal{N}(0, \sigma_w^2)$: zero mean white noise

$s_k \sim \mathcal{N}(0, \sigma_s^2)$: zero mean white noise

w_k, s_k uncorrelated

Now

$$\mathbf{R}_0 = \sigma_w^2 \mathbf{I}, \quad \mathbf{R}_1 = (\sigma_s^2 + \sigma_w^2)\mathbf{I}$$

and

$$\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_w^2} \frac{1}{\sigma_w^2} \mathbf{I}$$

Hence, defining

$$\text{SNR} = \sigma_s^2 / \sigma_w^2$$

c_i is the constant

$$\begin{aligned}
c_i &= \sigma_w^2/(\sigma_w^2+\sigma_s^2) = 1/(1+\mathrm{SNR}) \\
1-c_i &= \mathrm{SNR}/(1+\mathrm{SNR}) \\
1/c_i - 1 &= 1/\mathrm{SNR} \\
\rho &= \overline{(1-c)}/\overline{(1/c-1)} = 1/(1+\mathrm{SNR})
\end{aligned}$$

Note the SNR is defined differently in the case of a zero mean stochastic signal and a non-zero mean deterministic signal (128).

INTERPRETATION: $1-c_i$ is the temporal "coherency function" (SNR normalized to interval $[0,1]$)) of the signal w.r.t. the measurement

$$\kappa := \sigma_s^2/(\sigma_w^2+\sigma_s^2) = \mathrm{SNR}/(1+\mathrm{SNR})$$

Thus LRT reduces to

$$T(\underline{x}) = \frac{\kappa}{\sigma_w^2}\sum_{k=1}^n x_k^2 \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

Which reduces to the Chi-squared test ("energy detector")

$$T'(\underline{x}) = \sum_{k=1}^n x_k^2/\sigma_w^2 \overset{H_1}{\underset{H_0}{\gtrless}} \chi_n^{-1}(1-\alpha)$$

NOTE: relation between Chi-square-mixture and Chi-square CDF's when $1-c_i =$ constant

$$\overline{\chi}_{n,\overline{(1-c)}} = n^{-1}\chi_n$$

Power curve reduces to

$$\beta = 1-\chi_n\left(\frac{1}{1+\mathrm{SNR}}\,\chi_n^{-1}(1-\alpha)\right)$$

Example 48 *Detection of uncorrelated non-stationary signal in noise*

$$\begin{aligned}
H_0: x_k &= w_k \\
H_1: x_k &= s_k + w_k
\end{aligned}$$

$w_k \sim \mathcal{N}(0,\sigma_w^2(k))$: uncorrelated noise samples

$s_k \sim \mathcal{N}(0,\sigma_s^2(k))$: uncorrelated signal samples

w_k, s_k uncorrelated

In this case

$$\mathbf{R}_0 = \text{diag}(\sigma_w^2(i)), \quad \mathbf{R}_1 = \text{diag}(\sigma_s^2(i) + \sigma_w^2(i))$$

and

$$\begin{aligned}\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1} &= \text{diag}\left(\frac{\sigma_s^2(i)}{\sigma_s^2(i)+\sigma_w^2(i)} \frac{1}{\sigma_w^2(i)}\right) \\ &= \text{diag}\left(\kappa_i/\sigma_w^2(i)\right)\end{aligned}$$

where κ_i is time varying coherency function

$$\kappa_i = \frac{\sigma_s^2(i)}{\sigma_s^2(i)+\sigma_w^2(i)}$$

is

Hence, MP-LRT of level α reduces to

$$\frac{1}{\overline{\kappa}} \sum_{k=1}^n \kappa_k \frac{x_k^2}{\sigma_w^2(k)} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad = \overline{\chi}_{n,\kappa}^{-1}(1-\alpha)$$

or equivalently in terms of the original $T(\underline{x})$

$$T(\underline{x}) = \sum_{k=1}^n \kappa_k \frac{x_k^2}{\sigma_w^2(k)} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma = \overline{\kappa}\, \overline{\chi}_{n,\kappa}^{-1}(1-\alpha)$$

Special case of white noise: $\sigma_w^2(k) = N_o/2$

$$\sum_{k=1}^n \kappa_k\, x_k^2 \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma = \frac{N_o}{2}\, \overline{\kappa}\, \overline{\chi}_{n,\kappa}^{-1}(1-\alpha)$$

TWO USEFUL INTERPRETATIONS

Assume white noise for simplicity (we know that we can simply prewhiten by $1/\sigma_w(k)$ if non-white w_k).

1. "MEMORYLESS" ESTIMATOR CORRELATOR IMPLEMENTATION

Rewrite test statistic as

$$\sum_{k=1}^n \hat{s}_k\, x_k \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

where $\hat{s}_k$ is LLMSE estimator of s_k given x_k

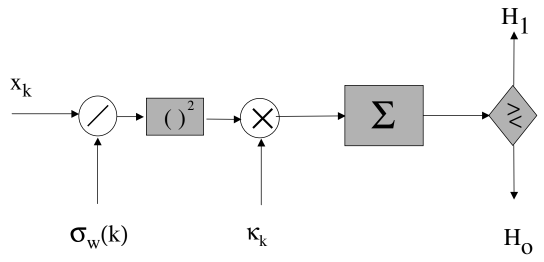


Figure 168: *LRT for detecting independent zero mean non-stationary Gaussian signal in non-stationary Gaussian noise.*

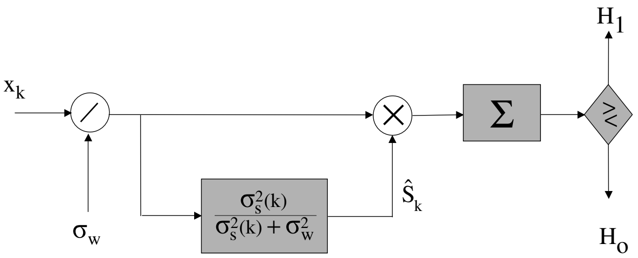


Figure 169: *Memoryless estimator correlator implementation of LRT for non-stationary uncorrelated signal in white noise. Note prewhitening operation* $1/\sigma_w^2$ *preceeds the estimator correlator.*

$$\hat{s}_k = \frac{\sigma_s^2(k)}{\sigma_s^2(k) + \sigma_w^2}\, x_k = \kappa_k\, x_k$$

2. "MEMORYLESS" FILTER-SQUARER IMPLEMENTATION

Rewrite test statistic as

$$\sum_{k=1}^{n} y_k^2 \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

where y_k is defined as

$$y_k = \sqrt{\frac{\sigma_s^2(k)}{\sigma_s^2(k) + \sigma_w^2}}\, x_k = \sqrt{\kappa_i}\, x_k$$

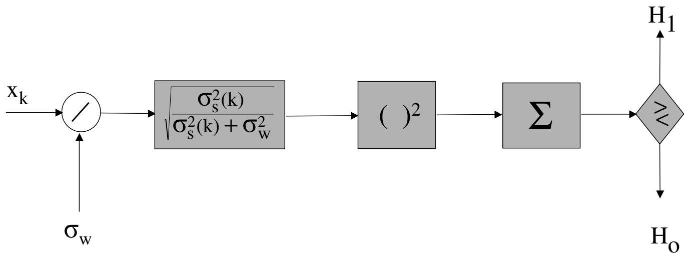


Figure 170: *Memoryless filter squarer implementation of LRT for non-stationary uncorrelated signal in white noise.*

POWER OF MEMORYLESS ESTIMATOR CORRELATOR:

as above

$$\beta = 1 - \overline{\chi}_{n,1/c-1}(\rho \overline{\chi}_{n,1-c}^{-1}(1-\alpha))$$

where

$$c_i = \frac{\sigma_w^2}{(\sigma_s^2(i) + \sigma_w^2)} = 1 - \kappa_i$$

$$\rho = \frac{\sum_i \kappa_i}{\sum_i \kappa_i/(1-\kappa_i)}$$

To a good approximation, at high SNR it can be shown that ROC depends on n, N_o, $\sigma_s^2(i)$ only through the following three SNR moments

$$\begin{aligned} SNR^{(1)} &= \frac{1}{\sigma_w^2}\sum_{i=1}^n \sigma_s^2(i) \\ SNR^{(2)} &= \left(\frac{1}{\sigma_w^2}\right)^2 \sum_{i,j=1}^n \sigma_s^2(i)\sigma_s^2(j) \\ SNR^{(3)} &= \left(\frac{1}{\sigma_w^2}\right)^3 \sum_{i,j,k=1}^n \sigma_s^2(i)\sigma_s^2(j)\sigma_s^2(k) \end{aligned}$$

Example 49 *Offline detection of w.s.s. signal in white noise*

Assume a window of n samples of a zero mean w.s.s. process x_k are available to test

$$\begin{aligned} H_0 : x_k &= w_k, \;\; k = 0, \ldots, n-1 \\ H_1 : x_k &= s_k + w_k \;\; k = 0, \ldots, n-1 \end{aligned}$$

where

* w_k: Gaussian white noise with PSD $\mathcal{P}_w(\omega) = N_o/2$

* s_k: zero mean w.s.s. Gaussian signal with known autocorrelation function $r_s(k) = E[s_l s_{l-k}]$

* w_k, s_k uncorrelated

The $n \times n$ covariance matrices under H_0 and H_1 are

$$\mathbf{R}_0 = \mathbf{R}_w = \sigma_w^2 \mathbf{I}, \quad \mathbf{R}_1 = \mathbf{R}_s + \sigma_w^2 \mathbf{I}$$

where $\mathbf{R}_s = ((\; r_s(l-m)\;))_{l,m=1}^n$ is an $n \times n$ p.d. Toeplitz matrix and $\sigma_w^2 = N_o/2$.

We know that MP-LRT is of form

$$T(\underline{x}) = \underline{x}^T(\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1})\underline{x} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \eta$$

However, in this form, the detector is not implementable for large n due to the need for to perform the $\mathbf{R}_s$ matrix inversion. An alternative, for large n, is to invoke a "spectral decomposition" of the Toeplitz matrix $\mathbf{R}_s$, sometimes known as the Grenander representation, pursued below.

Define $\mathbf{E}$ the $n \times n$ unitary "DFT matrix" with n columns $\underline{e}_k$ given by

$$\underline{e}_k = [1, e^{j\omega_k}, \ldots, e^{j\omega_k(n-1)}]^T/\sqrt{n}$$

$j = \sqrt{-1}$, and $\omega_k = \frac{k}{n}2\pi \in [0, 2\pi)$, $k = 0, \ldots, n-1$, is the k-th radian frequency. Let $\underline{\tilde{x}} = [\tilde{x}_1, \ldots, \tilde{x}_n]^T$ denote the vector of ($\sqrt{n}$-normalized) DFT coefficients associated with $\underline{x} = [x_1, \ldots, x_n]^T$:

$$\underline{\tilde{x}} = E^H \underline{x}.$$

Then as $\mathbf{E}\mathbf{E}^H = \mathbf{I}$, we can represent the LRT statistic as

$$\begin{aligned} \underline{x}^T(\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1})\underline{x} &= [\mathbf{E}^H\underline{x}]^H(\mathbf{E}^H\mathbf{R}_0^{-1}\mathbf{E} - \mathbf{E}^H\mathbf{R}_1^{-1}\mathbf{E})[\mathbf{E}^H\underline{x}] \\ &= \underline{\tilde{x}}^H(\frac{1}{\sigma_w^2}\mathbf{I} - \mathbf{E}^H\mathbf{R}_1^{-1}\mathbf{E})\underline{\tilde{x}} \end{aligned} \tag{129}$$

Now, remarkably, for large n the matrix $\mathbf{E}^H\mathbf{R}_1^{-1}\mathbf{E}$ is approximately diagonal, i.e. the DFT operator $\mathbf{E}$ diagonalizes the covariance matrix. To show this we first state a theorem:

Spectral approximation Theorem [11]: For any positive definite Toeplitz matrix $\mathbf{R} = ((r_{i-j}))_{i,j=1}^n$

$$\mathbf{E}^H\mathbf{R}\mathbf{E} = \mathrm{diag}_k(\mathrm{DFT}_r(\omega_k)) + O(1/n)$$

where $\mathrm{DFT}_r(\omega_k) = \sum_{k=-n+1}^{n-1} r_k e^{-j\omega k}$ is the k-th coefficient of the discrete fourier transform (DFT) of the sequence $r = \{r_{-n+1}, \ldots, r_0, \ldots, r_{n-1}\}$, and $O(1/n)$ is a term that goes to zero at rate $1/n$. This implies that for large n the eigenvectors of $\mathbf{R}$ are the DFT vectors and the eigenvalues are the DFT coefficients of the distinct elements of $\mathbf{R}$.

Proof of spectral approximation theorem:

It suffices to show that as $n \to \infty$ the DFT matrix $\mathbf{E}$ asymptotically diagonalizes $\mathbf{R}$, i.e.,

$$\underline{e}_k^H\mathbf{R}\underline{e}_l \to \begin{cases} \mathrm{DFT}_r(\omega_k), & k = l \\ 0, & o.w. \end{cases}$$

So let's write out the quadratic form explicitly

$$\underline{e}_k^H\mathbf{R}\underline{e}_l \quad = n^{-1}\sum_{p=0}^{n-1}\sum_{m=0}^{n-1} e^{-j\omega_k p} e^{\omega_l m} r_{p-m}$$

Next we rearrange the summation a bit

$$\underline{e}_k^H\mathbf{R}\underline{e}_l \quad = n^{-1}\sum_{m=0}^{n-1} e^{j(\omega_l-\omega_k)m} \sum_{p=0}^{n-1} e^{-j\omega_k(p-m)} r_{p-m}$$

Now we make a change of indices in the summations $m \to t \in \{1, \ldots, n\}$ and $m - p \to \tau \in \{-n, \ldots, n\}$ to obtain

$$\underline{e}_k^H\mathbf{R}\underline{e}_l \quad = \sum_{\tau=-n+1}^{n-1} r_\tau e^{-j\omega_k(\tau)} \underbrace{n^{-1}\sum_{t=0}^{n-1} e^{j(\omega_l-\omega_k)t}}_{g_n(\omega_l-\omega_k)}$$

where $g_n(u) = n^{-1}(e^{jun/2}\sin(u(n+1)/2)/\sin(u/2) - 1)$. Now, as $n \to \infty$, the term $g_n(\omega_l - \omega_k)$ converges to a discrete delta function:

$$\lim_{n\to\infty} g_n(\omega_l - \omega_k) = \delta_{k-l}$$

and so, assuming appropriate conditions allowing us to bring the limit under the summation, we have the large n approximation

$$\underline{e}_k^H\mathbf{R}\underline{e}_l \quad = \sum_{\tau=-n+1}^{n-1} r_\tau e^{-j\omega_k\tau}\ \delta_{k-l} = \mathrm{DFT}_r(\omega_k)\delta_{k-l}$$

which establishes the spectral approximation. ◇

Applying the spectral decomposition theorem to the Toeplitz matrix $\mathbf{R}_1^{-1} = [\mathbf{R}_s + \sigma_w^2 \mathbf{I}]^{-1}$ (the inverse of a Toeplitz matrix is Toeplitz) we obtain

$$\mathbf{E}^H \mathbf{R}_1^{-1} \mathbf{E} = \text{diag}_k \left\{ \frac{1}{\mathcal{P}_s(\omega_k) + \sigma_w^2} \right\} + O(1/n)$$

where $\mathcal{P}_s(\omega_k)$ is the power spectral density associated with s_k, i.e., the DFT of $\{r_s(-n+1), \ldots, r_s(n-1)\}$. We have from (129) the following form of the MP-LRT test statistic (recall that $\sigma_w^2 = N_o/2$)

$$\begin{aligned} T(\underline{x}) &= \underline{\tilde{x}}^H (\frac{1}{\sigma_w^2}\mathbf{I} - \mathbf{E}^H \mathbf{R}_1^{-1} \mathbf{E}) \underline{\tilde{x}} & (130) \\ &= \frac{2}{N_o}\, \underline{\tilde{x}}^H \, \text{diag}\left(\kappa(\omega_k)\right)\, \underline{\tilde{x}} & (131) \end{aligned}$$

where $\kappa(\omega)$ is the spectral coherency function

$$\kappa(\omega) = \frac{\mathcal{P}_s(\omega)}{\mathcal{P}_s(\omega) + N_o/2}$$

Expressing the quadratic form as a sum we obtain the equivalent large n form for the MP-LRT

$$T(\underline{x}) = \frac{2}{N_o} \sum_{k=0}^{n-1} \frac{\mathcal{P}_s(\omega_k)}{\mathcal{P}_s(\omega_k) + N_o/2} \; |\underline{\tilde{x}}_k|^2 \; 1/n \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

where, as before, γ is the level α threshold

$$\gamma = \overline{\kappa}\, \overline{\chi}_{n,\kappa}(1-\alpha)$$

and $\{\sqrt{n}\tilde{x}_k\}$ are the DFT coefficients of the observations. The quantity $|\tilde{x}_k|^2 \; 1/n$ is known as the Periodogram estimate of the PSD of x_k.

IMPLEMENTATION ISSUES

Using the duality between convolution in the time domain and multiplication in the frequency domain, identify the test statistic as:

$$T(\underline{x}) = \frac{2}{N_o} \sum_{k=0}^{n-1} \underbrace{\frac{\mathcal{P}_s(\omega_k)}{\mathcal{P}_s(\omega_k) + N_o/2}}_{(\hat{S}(\omega_k))^*} \tilde{x}_k^* \; \tilde{x}_k = \frac{2}{N_o} \sum_{k=0}^{n-1} \hat{s}_k x_k,$$

where $\hat{s}_k$ is the inverse DFT of $\hat{S}(\omega_k)$.

Implementation 1: Estimator correlator:

Absorbing $N_o/2$ into the threshold, the MP-LRT can be written as

$$\sum_{k=0}^{n-1} \hat{s}_k x_k \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma = \frac{N_o}{2}\, \overline{\kappa}\, \overline{\chi}_{n,\kappa}(1-\alpha)$$

where

$$\hat{s}_k = h_{\text{MMSE}}(k) * x_k$$

and $h_{\text{MMSE}}(k)$ is the Wiener filter with transfer function

$$H_{\text{MMSE}}(\omega) = \frac{\mathcal{P}_s(\omega)}{\mathcal{P}_s(\omega) + N_o/2}$$

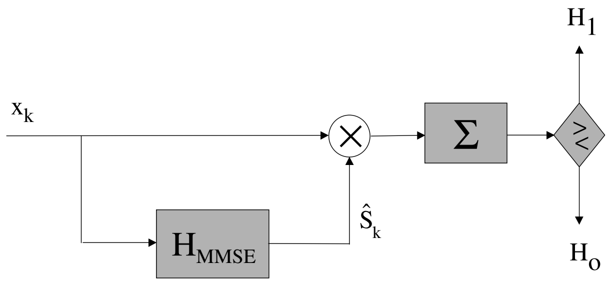


Figure 171: *Estimator correlator implementation of LRT for w.s.s. signal in white noise.*

Alternatively use

Parseval's theorem: if $f(k) \Leftrightarrow F(\omega_k)$ are DFT pair then

$$n^{-1} \sum_{k=-\infty}^{\infty} |F(\omega_k)|^2 = \sum_{k=0}^{n-1} |f(k)|^2$$

Implementation 2: filter-squarer

$$\sum_{k=0}^{n-1} y_k^2 \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma = \frac{N_o}{2} \, \overline{\kappa} \, \overline{\chi}_{n,\kappa}(1-\alpha)$$

where

$$y_k = h_k * x_k$$

and h_k has transfer function

$$H(\omega) = \sqrt{\frac{\mathcal{P}_s(\omega)}{\mathcal{P}_s(\omega) + N_o/2}}$$

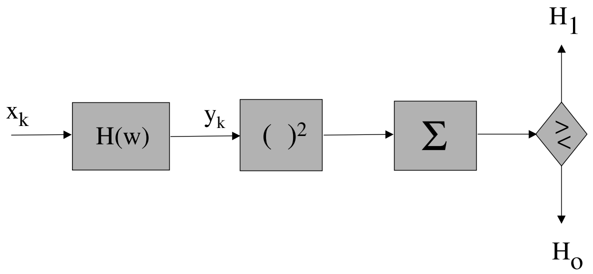


Figure 172: *Filter squarer implementation of LRT for w.s.s. signal in white noise.*

ROC: identically to previous example

$$\beta = 1 - \overline{\chi}_{n,1/c-1}(\rho \overline{\chi}^{-1}_{n,1-c}(1-\alpha))$$

except now $c = [c_1, \ldots, c_n]$ is

$$c_i = (N_o/2)/(\mathcal{P}_s(\omega_i) + N_o/2)$$

11.3 DETECTION OF NON-ZERO MEAN NON-STATIONARY SIGNAL IN WHITE NOISE

Now consider

$$\begin{aligned} H_0 : x_k &= w_k \\ H_1 : x_k &= s_k + w_k \end{aligned}$$

* $w_k \sim \mathcal{N}(0, \sigma_w^2)$: white noise
* $s_k \sim \mathcal{N}(\mu_k, \sigma_s^2(k))$: uncorrelated signal samples
* w_k, s_k uncorrelated

Recall general formula for nonequal means and covariances for LRT

$$T(\underline{x}) = \tfrac{1}{2}\underline{x}^T[\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}]\underline{x} + (\underline{\mu}_1^T \mathbf{R}_1^{-1} - \underline{\mu}_0^T \mathbf{R}_0^{-1})\underline{x} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

For present case $\underline{\mu}_0 = 0$, $\mathbf{R}_1$ and $\mathbf{R}_0$ are diagonal and LRT statistic reduces to

$$\begin{aligned} T(\underline{x}) &= \frac{1}{2\sigma_w^2} \sum_{k=1}^{n} \frac{\sigma_s^2(k)}{\sigma_s^2(k)+\sigma_w^2}\, x_k^2 + \sum_{k=1}^{n} \frac{1}{\sigma_s^2(k)+\sigma_w^2}\, \mu_k x_k \\ &= \frac{1}{2\sigma_w^2} \left(\sum_{k=1}^{n} \kappa_k x_k^2 + 2\sum_{k=1}^{n} (1-\kappa_k)\mu_k x_k \right) \end{aligned}$$

It is easily shown (see exercises) that this LRT is equivalent to the test

$$\sum_{k=1}^{n} \frac{\sigma_s^2(k)}{\sigma_s^2(k)+\sigma_w^2}\, (x_k-\mu_k)^2 + 2\sum_{k=1}^{n} \mu_k x_k \underset{H_0}{\overset{H_1}{\gtrless}} \gamma' \qquad (132)$$

This test can be implemented by a combination of estimator-correlator and matched filter.

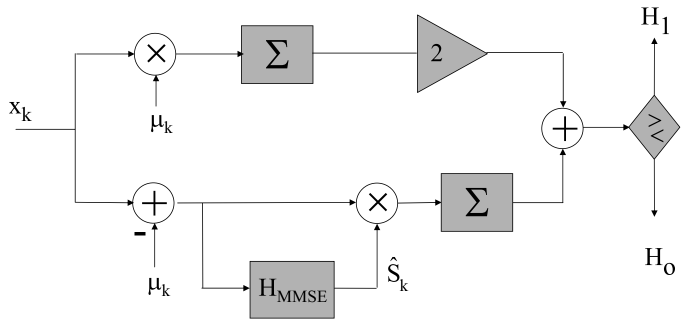


Figure 173: *Estimator correlator plus matched filter implementation of LRT for non-zero mean w.s.s. signal in white noise.*

PERFORMANCE:

The test statistic is now distributed as a noncentral Chi-square-mixture under H_0 and H_1 and analysis is somewhat more complicated (Johnson, Kotz and Balakrishnan [30, Sec. 18.8]).

11.4 ONLINE IMPLEMENTATIONS OF OPTIMAL DETECTORS

Objective: perform optimal detection at each sampling time $n = 1, 2, \ldots$ based only on past observations $0 < k \leq n$

$$
\begin{aligned}
H_0 : x_k &= v_k \\
H_1 : x_k &= s_k + v_k
\end{aligned}
$$

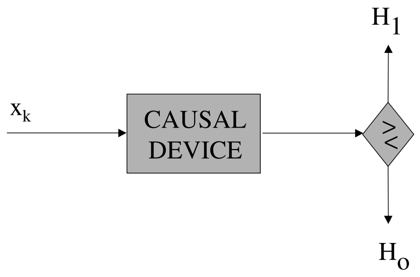


Figure 174: *Online detection seeks to develop optimal detector for each time instant n based only on past measurements.*

11.4.1 ONLINE DETECTION FOR NON-STATIONARY SIGNALS

Objective: decide between the presence of either of two random signals based on finite past $0 < k \le n$

$$
\begin{aligned}
H_0 : x_k &= s_0(k) + w_k \\
H_1 : x_k &= s_1(k) + w_k
\end{aligned}
$$

where

v_k: non-stationary zero mean Gaussian noise

$s_0(k),\ s_1(k)$: non-stationary zero mean Gaussian signals with known state space representations as in Sec. 6.7.1.

Recall: general MP-LRT is of form

$$T(\underline{x}) = \underline{x}^T [\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}] \underline{x} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \quad \gamma$$

Difficulty: growing memory in n makes computation of $T(x)$ impractical

Solution 1: Online dual Kalman signal selector

Solution 2: Online signal detector via Cholesky

11.4.2 ONLINE DUAL KALMAN SIGNAL SELECTOR

Let $\underline{\eta}_0$ and $\underline{\eta}_1$ denote vectors of innovations generated by Kalman filters matched to H_0 and H_1, respectively (See Sec. 6.8.2 to brush up on Kalman filters).

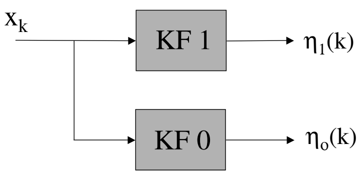


Figure 175: *Dual Kalman filters generate innovations processes η_1 and η_0*

We know that

$$\underline{\eta}_0 = \mathcal{A}_0 \underline{x}, \qquad \underline{\eta}_1 = \mathcal{A}_1 \underline{x}$$

$$\mathbf{R}_0 = \mathcal{A}_0^{-1} \mathbf{R}_{\eta 0} \mathcal{A}_0^{-T}, \qquad \mathbf{R}_1 = \mathcal{A}_1^{-1} \mathbf{R}_{\eta 1} \mathcal{A}_1^{-T}$$

where

* $\mathcal{A}_0, \mathcal{A}_1$ are lower triangular matrices of prediction coefficients

* $\mathbf{R}_{\eta 0}, \mathbf{R}_{\eta 1}$ are diagonal matrices of prediction error variances

Recall important property of innovations:

$\eta(k) = x(k) - \hat{x}(k|k-1) = x(k) - \hat{s}(k|k-1)$

* $E[\eta(k)] = 0$

* $\mathrm{var}(\eta(k)) = \underbrace{\underline{c}^T \mathbf{R}_{\tilde{\xi}}(k, k-1)\underline{c}}_{\sigma^2_{\tilde{s}}(k)} + \sigma_v^2$ is minimum prediction error variance

* $\left\{\frac{\eta(k)}{\sqrt{\mathrm{var}(\eta(k))}}\right\}$ is white

* $[\eta_1, \ldots, \eta_n]^T \sim \mathcal{N}_n(0, \mathrm{diag}(\mathrm{var}(\eta(k)))I)$

Using innovations representation we can re-express LR statistic

$$\begin{aligned} T(\underline{x}) &= \underline{x}^T[\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}]\underline{x} \\ &= \underline{x}^T[\ \mathcal{A}_0^T \mathbf{R}_{\eta_0}^{-1} \mathcal{A}_0 \ - \ \mathcal{A}_1^T \mathbf{R}_{\eta_1}^{-1} \mathcal{A}_1\ ]\underline{x} \\ &= [\mathcal{A}_0 \underline{x}]^T\ \mathbf{R}_{\eta_0}^{-1}\ \underbrace{[\mathcal{A}_0 \underline{x}]}_{\underline{\eta}_0} - [\mathcal{A}_1 \underline{x}]^T\ \mathbf{R}_{\eta_1}^{-1}\ \underbrace{[\mathcal{A}_1 \underline{x}]}_{\underline{\eta}_1} \end{aligned}$$

Or, LRT reduces to

$$T(\underline{x}) = \sum_{i=1}^n \frac{\eta_0^2(i)}{\mathrm{var}(\eta_0(i))} - \sum_{i=1}^n \frac{\eta_1^2(i)}{\mathrm{var}(\eta_1(i))} \ \overset{H_1}{\underset{H_0}{\gtrless}} \ \gamma$$

where, level α threshold is time varying. For example if $\mathbf{R}_0^{-1} > \mathbf{R}_1^{-1}$

$$\gamma = n\overline{(1-c)}\ \overline{\chi}^{-1}_{n,1-c}(1-\alpha)$$

Special Case: SIGNAL DETECTION IN WHITE NOISE

Here $s_0(k)$ is zero and v_k is white

$$\begin{aligned} H_0 : x_k &= v_k \\ H_1 : x_k &= s_1(k) + v_k \end{aligned}$$

and

* $\hat{s}_0(k|k-1) = 0$

* $\eta_0(k) = x_k$

* $\mathrm{var}_0(\eta_0(k)) = \sigma_v^2$

Thus MP-LRT simplifies to a “measured energy” vs. “Kalman residual” detector

$$T(\underline{x}) = \frac{1}{\sigma_v^2} \sum_{i=1}^n x_i^2 - \sum_{i=1}^n \frac{\eta_1^2(i)}{\mathrm{var}(\eta_1(i))} \ \overset{H_1}{\underset{H_0}{\gtrless}} \ \gamma$$

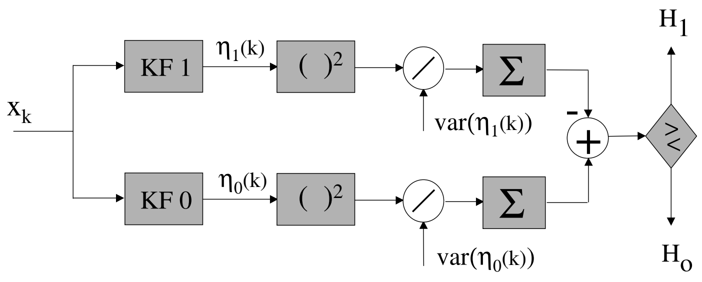


Figure 176: *Dual Kalman filter implementation of state space signal selector.*

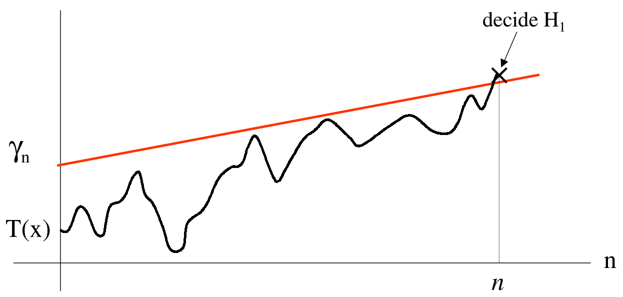


Figure 177: *Trajectory of dual Kalman filter implementation of state space signal selector. Note that the threshold is a function of time. If the number of samples n is random then the threshold of the test must be set by the method of repeated tests of significance.*

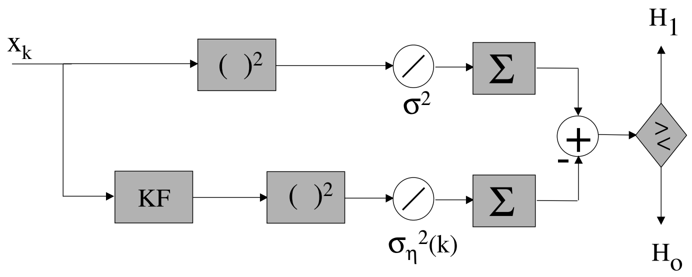


Figure 178: *The optimal detector of a single state space signal in noise.*

11.4.3 ONLINE SIGNAL DETECTOR VIA CHOLESKY

Again assume $s_0(k)$ is zero so that

$$\begin{aligned} H_0 : x_k &= v_k \\ H_1 : x_k &= s_1(k) + v_k \end{aligned}$$

Solution: apply Cholesky decomposition to $\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}$

Note

$$\begin{aligned} \mathbf{R}_0^{-1} - \mathbf{R}_1^{-1} &= \mathbf{R}_v^{-1} - [\mathbf{R}_s + \mathbf{R}_v]^{-1} \\ &= [\mathbf{R}_s + \mathbf{R}_v]^{-\frac{1}{2}} \mathbf{R}_s^{\frac{1}{2}} \mathbf{R}_v^{-1} \mathbf{R}_s^{\frac{1}{2}} [\mathbf{R}_s + \mathbf{R}_v]^{-\frac{1}{2}} \\ &> 0 \end{aligned}$$

Hence we can apply the Cholesky decomposition

$$\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1} = \mathbf{L}^T \, \mathbf{P} \, \mathbf{L}$$

* $\mathbf{L}$ is lower triangular matrix of “backward predictor coefficients”

* **P** is diagonal matrix of "backward predictor error variances"

Now apply Cholesky decomposition to $T(\underline{x})$

$$\begin{aligned} T(\underline{x}) &= \tfrac{1}{2}\underline{x}^T[\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}]\underline{x} \\ &= \tfrac{1}{2}\underline{x}^T[\mathbf{L}^T\ \mathbf{P}\ \mathbf{L}]\underline{x} \\ &= \tfrac{1}{2}[\mathbf{L}\underline{x}]^T\ \mathbf{P}\ \underbrace{[\mathbf{L}\underline{x}]}_{\underline{y}} \end{aligned}$$

or we have representation

$$T(\underline{x}) = \tfrac{1}{2}\sum_{i=1}^n \sigma_i^2\ y_i^2$$

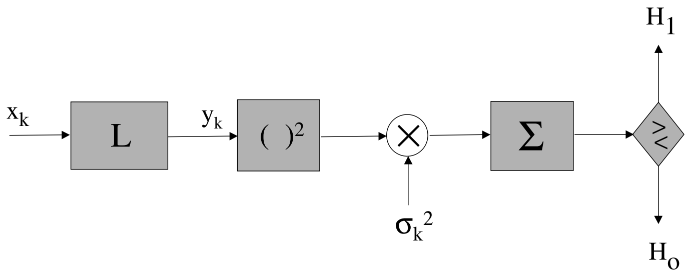


Figure 179: *On-line implementation of non-stationary signal detector via using Cholesky factor* **L**.

The MP level α test is simply

$$\sum_{i=1}^n \sigma_i^2\ y_i^2 \ \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array}\ \gamma_n = n\overline{(1-c)}\ \overline{\chi}_{n,\overline{\kappa}}^{-1}(1-\alpha)$$

where σ_i^2 and $\kappa = n\overline{(1-c)}$ are computed offline. Note that while we can generate a recursion for the test statistic, a recursion for the threshold γ_n is not available.

In many cases y_i can be generated by a "lumped" Kalman filter matched to a state observation model

$$
\begin{aligned}
x_k^{'} &= s_k^{'} + v_k^{'} \\
s_k^{'} &= \underline{c}_k^T \underline{\nu}_k \\
\underline{\nu}_{k+1} &= \mathbf{D}_k \underline{\nu}_k + \mathbf{E}_k \underline{w}_k^{'}
\end{aligned}
$$

synthesized such that the measurement covariance satisfies

$$\mathbf{R}_{x'} = \mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}$$

11.5 STEADY-STATE STATE-SPACE SIGNAL DETECTOR

Assume:

* State model for s_1 is LTI

* measurement noise v_k is w.s.s.

* limiting state error covariance matrix $\mathbf{R}_{\tilde{\xi}}(\infty)$ is non-singular

* Kalman filter is in steady state (n large)

Then, as innovations are w.s.s., the MP-LRT statistic can be written

$$T(\underline{x}) = \frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 \; - \; \frac{1}{2\mathrm{var}(\eta_1)} \sum_{i=1}^n \eta_1^2(i) \quad \begin{smallmatrix} H_1 \\ > \\ < \\ H_0 \end{smallmatrix} \quad \gamma$$

Or, using asymptotic innovations variance, we have MP test

$$\sum_{i=1}^n x_i^2 \; - \; \frac{\sigma_v^2}{\sigma_{\tilde{s}}^2 + \sigma_v^2} \sum_{i=1}^n \eta_1^2(i) \quad \begin{smallmatrix} H_1 \\ > \\ < \\ H_0 \end{smallmatrix} \quad \gamma$$

APPROXIMATE GLRT FOR UNKNOWN MEASUREMENT NOISE VARIANCE

We can implement an approximate GLRT to handle the case where the variance of the observation noise σ_v^2 is unknown. For this case the GLR statistic is

$$\Lambda_{\mathrm{GLR}} = \frac{\max_{\sigma_v^2>0} \, (\sigma_{\tilde{s}}^2 + \sigma_v^2)^{-n/2} \exp\left(-\frac{1}{2(\sigma_{\tilde{s}}^2+\sigma_v^2)} \, \sum_{i=1}^n \eta_1^2(i)\right)}{\max_{\sigma_v^2>0} \, (\sigma_v^2)^{-n/2} \exp\left(-\frac{1}{2\sigma_v^2} \, \sum_{i=1}^n x_i^2\right)}$$

Of course the denominator is maximized for

$$\sigma_v^2 = n^{-1} \sum_{i=1}^{n} x_i^2$$

and, as for the numerator, we proceed by an iterative approximation. First neglect the dependence of of η_1 on σ_v^2. Then the numerator is maximized for

$$\hat{\sigma_v}^2(n) = n^{-1} \sum_{i=1}^{n} \eta_1^2(i) - \sigma_s^2$$

Now generate $\eta_1(n+1)$ from the Kalman Filter having parameters A, b, c and $\hat{\sigma_v}^2$. In this way we obtain an approximate GLRT which is implemented by comparing the ratio of two variance estimators to a threshold. Note that the numerator and denominator of the test statistic are dependent so this is not an F-test.

$$\frac{\hat{\sigma}_v^2}{\hat{\sigma}_{\eta_1}^2} = \frac{\sum_{i=1}^n x^2(i)}{\sum_{i=1}^n \eta_1^2(i)} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

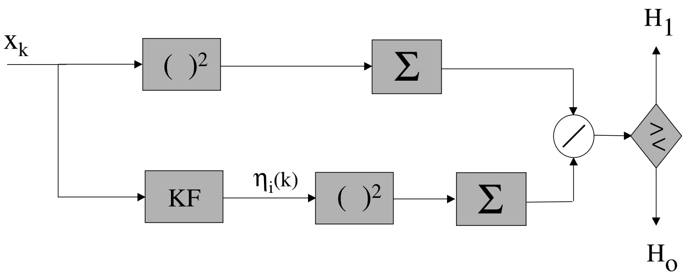


Figure 180: *Approximate steady-state GLRT signal detector for unknown measurement noise*

In analogous manner, for the GLRT signal selector we obtain

$$\frac{\sum_{i=1}^n \eta_0^2(i)}{\sum_{i=1}^n \eta_1^2(i)} \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

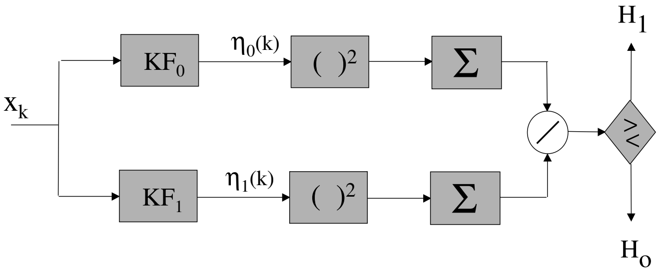


Figure 181: *GLRT signal selector for unknown measurement noise*

11.6 BACKGROUND REFERENCES

A concise mathematical statistics development of binary hypothesis testing for multivariate Gaussian observations can be found in Morrison [50]. For a signal detection perspective the books by Van Trees volume I [73] and volume III [74], and Whalen [76] are classics in the field. Other more recent signal processing oriented textbooks with coverage of this topic are [25], Poor [55], and Srinath, Rajasekaran and Viswanath [67]. A discussion of online implementations of optimal detection for random processes is treated in the context of change detection in the edited book by Basseville and Benveniste [3].

11.7 EXERCISES

11.1 Let $\underline{x} = [x_1, \ldots, x_n]^T$ be n samples of a waveform. It is of interest to test the two hypotheses

$$\begin{aligned} H_0 : \underline{x} &= a\underline{y} + \underline{w} \\ H_1 : \underline{x} &= \underline{s} + a\underline{y} + \underline{w} \end{aligned}$$

where $\underline{w}$ is zero mean Gaussian white noise, $\mathrm{cov}(\underline{w}) = \sigma^2 I$, $\underline{s}$ and $\underline{y}$ are known waveforms, and the scalar constant a is unknown.

(a) Assuming that a is a Gaussian r.v. with zero mean and variance σ_a^2 derive the MP LRT (with threshold) to test H_0 vs. H_1. Assume that a is independent of $\underline{w}$. Is this a UMP test for the case that the signal shape $\underline{s}/\|\underline{s}\|$ is known but its energy $\|\underline{s}\|^2$ is unknown? How about when signal shape $\underline{y}/\|\underline{y}\|$ is known but $\|\underline{y}\|^2$ is unknown?

(b) Under the assumption on a of part (a) find the detectibility index d which controls the ROC curve. Assume that $\|\underline{s}\| \leq 1$. Show that the ROC curve is optimized (maximum d) when the signal $\underline{s}$ is orthogonal to the interferer $\underline{y}$ but is otherwise arbitrary (Hint: you might want to use the Woodbury matrix identity).

(c) Assuming that a is a deterministic unknown constant, repeat parts (a) and (b) for the GLRT of H_0 vs. H_1.

11.2 Let x_k, $k = 1, \ldots, n$ be a segment of a discrete time random process. It is desired to test whether x_k contains a harmonic component (sinusoidal signal) or not

$$
\begin{aligned}
H_0 : x_k &= w_k \\
H_1 : x_k &= A\cos(\omega_o k + \psi) + w_k
\end{aligned}
$$

where w_k is zero mean Gaussian white noise with acf $r_w(k) = N_0/2\delta_k$, $\omega_o = 2\pi l/n$ for some integer l, A is a deterministic amplitude, and ψ is a uniform phase over $[0, 2\pi]$. The random phase of the sinusoid and the noise samples are independent of each other.

(a) Show that under H_1 the auto-correlation function of x_k is $E[x_i x_{i-k}] = r_x(k) = A^2/2\cos(\omega_o k) + N_0/2\delta_k$ and derive the PSD $\mathcal{P}_x$.

(b) Derive the MP LRT with threshold and implement the MP LRT as an estimator correlator and a filter squarer. (Hint: as ψ is uniform and $f_1(\underline{x}|\psi)$ is a Gaussian p.d.f. $f_1(\underline{x}) = (2\pi)^{-1}\int_0^{2\pi} f_1(\underline{x}|\psi)d\psi$ is a Bessel function of the form $B_0(r) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{r\cos\psi}d\psi$ which is monotone in a test statistic which under H_0 is distributed as a Chi-square with 2 df, i.e exponential.)

(c) Show that the MP LRT can also be implemented as a test on the *periodogram spectral estimator* $\mathcal{P}_{per}(\omega_o) = \frac{1}{n}|DFT\{x_k\}_{\omega=\omega_o}|^2$ where $DFT\{x_k\}_\omega = \sum_{k=1}^n x_k e^{-j\omega k}$ is the DTFT of $\{x_k\}_{k=1}^n$, $\omega \in \{2\pi l/n\}_{l=1}^n$.

11.3 Find the GLRT for the previous problem under the assumption that both A and ω_o are unknown (Hint: as no closed form solution exists for the MLE's of A and ω_o you can leave your answer in the form of a "peak detector" block diagram).

11.4 Derive the "completion of the square" result (Eq. (132) in section 11.3.

11.5 A sensor is placed on a North Atlantic oil derick at a particular spatial location to monitor the mechanical state of the structure. When the mechanical state is "normal" the sensor produces a measurement which follows the state space model:

$$
\begin{aligned}
x_k &= s_k + v_k \\
s_{k+1} &= a s_k + w_k
\end{aligned}
$$

$k = 0, 1, \ldots$. A model for impending failure of the mechanical structure is that a shift in the damping constant a occurs. Assuming the standard Gaussian assumptions on the dynamical model under both normal and failure modes, the detection of impending failure can be formulated as testing between

$$
\begin{aligned}
H_0 : a &= a_o \\
H_1 : a &\neq a_o
\end{aligned}
$$

where $a_o \in (-1, 1)$ is known.

(a) Implement the MP test of level α for the simple alternative $H_1 : a = a_1$, where $a_1 \neq a_0$, with a pair of Kalman filters. If you solved Exercise 6.14 give explicit forms for your filters using the results of that exercise.

(b) Now treat the general composite case above with your favorite method, e.g. LMP or GLRT. Take this problem as far as you can by making simplifying assumptions starting with assuming steady state operation of the Kalman filters.

11.6 Available for observation are n time samples $X(k)$,

$$X(k) = \sum_{i=1}^{p} \alpha_i g_i(k - \tau_i) + W(k), \quad k = 1, \ldots, n$$

where $W(k)$ is a zero mean Gaussian white noise with variance $\mathrm{var}(W(k)) = \sigma_w^2$, α_i, $i = 1, \ldots, p$, are p i.i.d. zero mean Gaussian random variables with variance σ_a^2, and $g_i(u)$, $i = 1, \ldots, p$, are p known time functions over $u \in (-\infty, \infty)$. The α_i and $W(k)$ are uncorrelated and p is known. Define K as the $p \times p$ matrix of inner products of the g_i's, i.e. K has entries $\kappa_{ij} = \sum_{k=1}^{n} g_i(k - \tau_i) g_j(k - \tau_j)$.

(a) Show that the ML estimator of the τ_i's involves maximizing a quadratic form $\underline{y}^T[I + \rho K]^{-1}\underline{y} - b$ where $\underline{y} = [y_1, \ldots, y_p]^T$ is a vector of p correlator outputs $y_i(\tau_i) = \sum_{k=1}^{n} x(k) g_i(k - \tau_i)$, $i = 1, \ldots, p$, $b = b(\underline{\tau})$ is an observation independent bias term, and $\rho = \sigma_a^2/\sigma_w^2$ is the SNR (Hint: express log-likelihood function in vector-matrix form and use a matrix inverse (Woodbury) identity). Draw a block diagram of your ML estimator implemented as a peak picker, i.e. a variable filter applied to the data over which you seek to maximize the output.

(b) Now consider the detection problem

$$\begin{aligned} H_0 &: X(k) = W(k) \\ H_1 &: X(k) = \sum_{i=1}^{p} \alpha_i g_i(k - \tau_i) + W(k) \end{aligned}$$

For known τ_i's derive the LRT and draw a block diagrom of the detector. Is the LRT UMP for unknown τ_i's? How about for known τ_i's but unknown SNR σ_a^2/σ_w^2?

(c) Now assume that the τ_i's are unknown and that the α_i's are also unknown and non-random. Show that in the GLRT the maximization over the α_i's can be performed explicitly. Draw a block diagram of the GLRT implemented with a thresholded peak picker over the τ_i's.

11.7 Observed is a random process $\{x_i\}_{i=1}^{n}$ consisting of Gaussian random variables. Assume that

$$x_k = s_k + w_k$$

where s_k and w_k are uncorrelated Gaussian variables with variances $\sigma_s^2(k)$ and σ_w^2, respectively. The noise w_k is white and s_k is uncorrelated over time but non-stationary, i.e., it has time varying variance. In this problem we assume that the instantaneous SNR $\gamma(k) = \sigma_s^2(k)/\sigma_w^2$ is known for all time but that the noise power level σ_w^2 is unknown.

(a) For known σ_w^2 and zero mean w_k and s_k derive the MP test of the hypotheses (no need to set the threshold)

$$\begin{aligned} H_0 &: x_k = w_k \\ & \qquad\qquad\qquad k = 1, \ldots, n \\ H_1 &: x_k = s_k + w_k \end{aligned}$$

Does there exist a UMP test for unknown σ_w^2? If so what is it?

(b) Find the GLRT for the above hypotheses for unknown σ_w^2 (no need to set the threshold).

(c) Now assume that s_k has non-zero but constant mean $\mu = E[s_k]$. Find the GLRT for unknown μ and σ_w^2 (no need to set the threshold).

11.8 Observed is a random process $\{x_i\}_{i=1}^n$ consisting of Gaussian random variables. Assume that

$$x_k = s_k + w_k$$

where s_k and w_k are zero mean uncorrelated Gaussian variables with variances $a^2\sigma_s^2(k)$ and σ_w^2, respectively. The noise w_k is white and s_k is uncorrelated over time but non-stationary, i.e., it has time varying variance.

(a) For known a^2, σ_s^2 and σ_w^2 derive the MP test of the hypotheses

$$\begin{array}{lcl} H_0 & : & x_k = w_k \\ & & \qquad\qquad\qquad k = 1, \ldots, n \\ H_1 & : & x_k = s_k + w_k \end{array}$$

You do not need to derive an expression for the threshold. Is your test UMP for unknown a^2? If not is there a condition on $\sigma_s^2(k)$ that would make your test UMP?

(b) Find the locally most powerful test for unknown $a^2 > 0$ for the above hypotheses. How does your test compare to the matched filter detector for detection of non-random signals?

(c) Now assume that s_k has non-zero but constant mean $\mu = E[s_k]$. Find the MP test. Is your test UMP for unknown $\mu \neq 0$ when all other parameters are known? If not find the GLRT for this case.

End of chapter

12 COMPOSITE HYPOTHESES IN THE MULTIVARIATE GAUSSIAN MODEL

In Chapter 9 we covered testing of composite hypotheses on the mean and variance for univariate i.i.d. Gaussian measurements. In Chapter 11 we covered simple hypotheses on the mean and covariance in the multivariate Gaussian distribution. In this chapter we extend the techniques developed in Chapters 9 and 11 to multivariate Gaussian measurements with composite hypotheses on mean and covariance. In signal processing this is often called the Gaussian multi-channel model as i.i.d. measurements are made of a Gaussian random vector, and each element of the vector corresponds to a separate measurement channel (see Fig. 182).

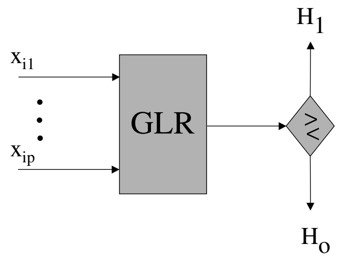


Figure 182: *GLR detector from multi-channel Gaussian measurements.*

Specifically, we will cover the following

* Double sided GLRT for equality of vector mean

* Double sided GLRT for equality two vector means

* Double sided GLRT for independence of samples

* Double sided GLRT for whiteness of samples

* Confidence regions for vector mean

Here the measurements are a set of n i.i.d. p-dimensional Gaussian vectors, each having mean vector $\underline{\mu}$ and $p \times p$ covariance $\mathbf{R}$:

$$\underline{X}_i = \begin{bmatrix} X_{i1} \\ \vdots \\ X_{ip} \end{bmatrix}, \quad i = 1, \ldots n$$

For notational convenience we denote the measurements by a random $p \times n$ *measurement matrix*

$$\mathbf{X} = [\underline{X}_1, \ldots, \underline{X}_n]$$

This matrix has the following properties:

* $\{\underline{X}_i\}_{i=1}^n$: independent Gaussian columns ($n \geq p$)

* $\underline{\mu} = E_\theta[\underline{X}_i]$: mean vector

* $\mathbf{R} = \mathrm{cov}_\theta(\underline{X}_i)$: covariance matrix ($p \times p$)

12.1 MULTIVARIATE GAUSSIAN MATRICES

In Section 3.1.1 of Chapter 3 we introduced the multivariate Gaussian density for random vectors. This is easily extended to the present case of random matrices $\mathbf{X}$ composed of i.i.d. columns of Gaussian random vectors. The jpdf of such a Gaussian matrix $\mathbf{X}$ has the form

$$\begin{aligned}
&f(\mathbf{X};\ \underline{\mu}, \mathbf{R}) \\
&= \left(\frac{1}{(2\pi)^p\ |\mathbf{R}|}\right)^{n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^n (\underline{X}_i - \underline{\mu})^T \mathbf{R}^{-1} (\underline{X}_i - \underline{\mu})\right) \\
&= \left(\frac{1}{(2\pi)^p\ |\mathbf{R}|}\right)^{n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^n \mathrm{trace}\left\{(\underline{X}_i - \underline{\mu})(\underline{X}_i - \underline{\mu})^T \mathbf{R}^{-1}\right\}\right)
\end{aligned}$$

This density can also be represented in more compact form as:

$$f(\mathbf{X};\ \underline{\mu}, \mathbf{R}) = \left(\frac{1}{(2\pi)^p\ |\mathbf{R}|}\right)^{n/2} \exp\left(-\frac{n}{2}\,\mathrm{trace}\{\hat{\mathbf{R}}_\mu\ \mathbf{R}\}\right)$$

where we have defined the $p \times p$ covariance estimator

$$\begin{aligned}
\hat{\mathbf{R}}_{\underline{\mu}} &= n^{-1}\sum_{i=1}^n (\underline{X}_i - \underline{\mu})(\underline{X}_i - \underline{\mu})^T \\
&= \frac{1}{n}(\mathbf{X} - \underline{\mu}\underline{1}^T)(\mathbf{X} - \underline{\mu}\underline{1}^T)^T.
\end{aligned}$$

12.2 DOUBLE SIDED TEST OF VECTOR MEAN

We pose the two hypotheses:

$$H_0 : \underline{\mu} = \underline{\mu}_o, \quad \mathbf{R} > 0 \tag{133}$$

$$H_1 : \underline{\mu} \neq \underline{\mu}_o, \quad \mathbf{R} > 0. \tag{134}$$

the GLRT of these hypotheses is

$$\Lambda_{\mathrm{GLR}} = \frac{\max_{\underline{\mu}, \mathbf{R}>0} f(\mathbf{X}; \underline{\mu}, \mathbf{R})}{\max_{\mathbf{R}>0} f(\mathbf{X}; \underline{\mu}_o, \mathbf{R})}.$$

Now, it is easily seen that

$$\max_{\underline{\mu},\mathbf{R}>0} f(\mathbf{X};\underline{\mu},\mathbf{R}) = \max_{\mathbf{R}>0} f(\mathbf{X};\ \underline{\overline{X}},\mathbf{R})$$

where the column sample mean is defined as

$$\underline{\overline{X}} = n^{-1}\sum_{i=1}^{n} \underline{X}_i = \mathbf{X}\underline{1}\,\frac{1}{n}.$$

Therefore, we can rewrite the GLRT as

$$\begin{aligned}
\Lambda_{\text{GLR}} &= \frac{\max_{\underline{\mu},\mathbf{R}>0} f(\mathbf{X};\underline{\mu},\mathbf{R})}{\max_{\mathbf{R}>0} f(\mathbf{X};\underline{\mu}_o,\mathbf{R})} \\
&= \frac{\max_{\mathbf{R}>0}\ |\mathbf{R}|^{-n/2}\ \exp\left(-\frac{1}{2}\text{ trace}\left\{\hat{\mathbf{R}}_{\underline{\overline{X}}}\mathbf{R}^{-1}\right\}\right)}{\max_{\mathbf{R}>0}\ |\mathbf{R}|^{-n/2}\ \exp\left(-\frac{n}{2}\text{ trace}\left\{\hat{\mathbf{R}}_{\underline{\mu}}\mathbf{R}^{-1}\right\}\right)}
\end{aligned}$$

FACT: for any vector $\underline{t} = [t_1,\ldots,p]^T$

$$\max_{\mathbf{R}>0}\ \left\{|\mathbf{R}|^{-n/2}\ \exp\left(-\frac{n}{2}\text{ trace}\left\{\hat{\mathbf{R}}_{\underline{t}}\mathbf{R}^{-1}\right\}\right)\right\} = |\hat{\mathbf{R}}_{\underline{t}}|^{-n/2}\ e^{-n^2/2}$$

and the maximum is attained by

$$\mathbf{R} = \hat{\mathbf{R}}_{\underline{t}} = n^{-1}\sum_{i=1}^{n}(\underline{X}_i - \underline{t})(\underline{X}_i - \underline{t})^T$$

Proof:

The maximizing $\mathbf{R}$ also maximizes

$$l(\mathbf{R}) = \ln f(\mathbf{X};\underline{t},\mathbf{R}) = \frac{n}{2}\ \ln|\mathbf{R}|\ -\frac{n}{2}\text{ trace}\left\{\hat{\mathbf{R}}_{\underline{t}}\mathbf{R}^{-1}\right\}$$

Define the transformed covariance $\tilde{\mathbf{R}}$

$$\tilde{\mathbf{R}} = \hat{\mathbf{R}}_{\underline{t}}^{-1/2}\ \mathbf{R}\ \hat{\mathbf{R}}_{\underline{t}}^{-1/2}.$$

Then, since the trace and the determinant satisfy

$$\text{trace}\{\mathbf{AB}\} = \text{trace}\{\mathbf{BA}\},\quad |\mathbf{AB}| = |\mathbf{BA}| = |\mathbf{B}|\ |\mathbf{A}|,$$

we have

$$\begin{aligned}
l(\mathbf{R}) &= -\frac{n}{2}\ \left(\ln|\hat{\mathbf{R}}_{\underline{t}}^{1/2}\tilde{\mathbf{R}}\hat{\mathbf{R}}_{\underline{t}}^{1/2}|\ +\ \text{trace}\left\{\hat{\mathbf{R}}_{\underline{t}}^{1/2}\tilde{\mathbf{R}}^{-1}\hat{\mathbf{R}}_{\underline{t}}^{-1/2}\right\}\right) \\
&= -\frac{n}{2}\ \left(\ln|\hat{\mathbf{R}}_{\underline{t}}\tilde{\mathbf{R}}|\ +\ \text{trace}\left\{\tilde{\mathbf{R}}^{-1}\right\}\right) \\
&= -\frac{n}{2}\ \left(\ln|\hat{\mathbf{R}}_{\underline{t}}| + \ln|\tilde{\mathbf{R}}|\ +\ \text{trace}\left\{\tilde{\mathbf{R}}^{-1}\right\}\right) \\
&= -\frac{n}{2}\ \left(\ln|\hat{\mathbf{R}}_{\underline{t}}| + \sum_{j=1}^{p}\ln\tilde{\lambda}_j + \sum_{j=1}^{p}\frac{1}{\tilde{\lambda}_j}\right)
\end{aligned}$$

where $\{\tilde{\lambda}_j\}$ are the eigenvalues of $\tilde{\mathbf{R}}$

Hence the maximizing $\mathbf{R}$ satisfies for $j = 1, \ldots, p$

$$\begin{aligned} 0 &= \frac{d}{d\tilde{\lambda}_j} l(\mathbf{R}) \\ &= -\frac{n}{2}\left(\frac{1}{\tilde{\lambda}_j} - \frac{1}{\tilde{\lambda}_j^2}\right) \end{aligned}$$

so that the maximizing $\tilde{\mathbf{R}}$ has identical eigenvalues

$$\tilde{\lambda}_j = 1, \quad j = 1, \ldots, p.$$

This implies that the maximizing $\tilde{\mathbf{R}}$ is an orthogonal (unitary) matrix $\mathbf{U}$. But, since $\tilde{\mathbf{R}}$ is also symmetric, $\tilde{\mathbf{R}}$ is in fact the $p \times p$ identity^o

Therefore

$$\mathbf{I} = \tilde{\mathbf{R}} = \hat{\mathbf{R}}_{\underline{t}}^{-1/2} \mathbf{R} \hat{\mathbf{R}}_{\underline{t}}^{-1/2}$$

giving the maximizing $\mathbf{R}$ as

$$\mathbf{R} = \hat{\mathbf{R}}_{\underline{t}},$$

as claimed. $\diamond$

Note: We have just shown that

1. The MLE of $\mathbf{R}$ for known $\underline{\mu} = \underline{\mu}_o$ is

$$\hat{\mathbf{R}}_{\underline{\mu}} = n^{-1} \sum_{i=1}^{n} (\underline{X}_i - \underline{\mu}_o)(\underline{X}_i - \underline{\mu}_o)^T.$$

2. The MLE of $\mathbf{R}$ for unknown $\underline{\mu}$ is

$$\hat{\mathbf{R}}_{\overline{\underline{X}}} = \hat{\mathbf{R}} = n^{-1} \sum_{i=1}^{n} (\underline{X}_i - \overline{\underline{X}})(\underline{X}_i - \overline{\underline{X}})^T.$$

Plugging the above MLE solutions back into GLRT statistic for testing (134)

$$\begin{aligned} \Lambda_{\text{GLR}} &= \left(\frac{|\hat{\mathbf{R}}_{\underline{\mu}_o}|}{|\hat{\mathbf{R}}|}\right)^{n/2} \\ &= \left(\left|\hat{\mathbf{R}}_{\underline{\mu}_o} \hat{\mathbf{R}}^{-1}\right|\right)^{n/2}. \end{aligned}$$

Using

$$\hat{\mathbf{R}}_{\underline{\mu}_o} = \hat{\mathbf{R}} + (\overline{\underline{X}} - \underline{\mu}_o)(\overline{\underline{X}} - \underline{\mu}_o)^T,$$

[0]If $\mathbf{U}$ is orthogonal then $\mathbf{U}^H = \mathbf{U}^{-1}$ matrix $\mathbf{I}$. If in addition $\mathbf{U}$ is symmetric then $\mathbf{U} = \mathbf{U}^T = \mathbf{U}^{-1}$ so that $\mathbf{U} = \mathbf{I}$.

we have the equivalent GLRT ($\Lambda_{\text{GLR}} = (T(\mathbf{X}))^{n/2}$)

$$\begin{aligned} T(\mathbf{X}) &= \left|\mathbf{I} + (\underline{\overline{X}} - \underline{\mu}_o)(\underline{\overline{X}} - \underline{\mu}_o)^T \hat{\mathbf{R}}^{-1}\right| \\ &= \left|\mathbf{I} + \underbrace{\hat{\mathbf{R}}^{-\frac{1}{2}}(\underline{\overline{X}} - \underline{\mu}_o)}_{\underline{u}} \underbrace{(\underline{\overline{X}} - \underline{\mu}_o)^T \hat{\mathbf{R}}^{-\frac{1}{2}}}_{\underline{u}^T}\right| \overset{H_1}{\underset{H_0}{\gtrless}} \gamma \end{aligned}$$

SIMPLIFICATION OF GLRT

Observe: $T(\mathbf{X})$ is the determinant of the sum of a rank 1 matrix and the identity matrix:

$$\begin{aligned} T(\mathbf{X}) &= \left|\mathbf{I} + \underbrace{\underline{u}\,\underline{u}^T}_{\text{rank} = 1}\right| \\ &= \prod_{j=1}^{p} \lambda_j \end{aligned}$$

where λ_j are the eigenvalues of the matrix $\mathbf{I} + \underline{u}\underline{u}^T$.

IMPORTANT FACTS:

1. Eigenvectors of $\mathbf{I} + \mathbf{A}$ are identical to eigenvectors of $\mathbf{A}$
2. Eigenvectors of $\mathbf{A} = \underline{u}\,\underline{u}^T$ are

$$\begin{aligned} \underline{\nu}_1 &= \underline{u}\frac{1}{\|\underline{u}\|} = \hat{\mathbf{R}}^{-1/2}(\underline{\overline{X}} - \underline{\mu}_o) \frac{1}{\sqrt{(\underline{\overline{X}} - \underline{\mu}_o)^T \hat{\mathbf{R}}^{-1}(\underline{\overline{X}} - \underline{\mu}_o)}} \\ \underline{\nu}_2, \ldots, \underline{\nu}_p &= \text{determined via Gramm-Schmidt.} \end{aligned}$$

3. Eigenvalues of $\mathbf{I} + \mathbf{A}$ are

$$\begin{aligned} \lambda_1 &= \underline{\nu}_1^T(\mathbf{I} + \mathbf{A})\underline{\nu}_1 = 1 + (\underline{\overline{X}} - \underline{\mu}_o)^T \hat{\mathbf{R}}^{-1}(\underline{\overline{X}} - \underline{\mu}_o) \\ \lambda_2 &= \ldots = \lambda_p = 1 \end{aligned}$$

Putting all of this together we obtain an equivalent expression for the GLRT of (134):

$$T(\mathbf{X}) = \prod_{j=1}^{p} \lambda_j = 1 + (\underline{\overline{X}} - \underline{\mu}_o)^T \hat{\mathbf{R}}^{-1}(\underline{\overline{X}} - \underline{\mu}_o) \overset{H_1}{\underset{H_0}{\gtrless}} \gamma$$

Or, equivalently, the GLRT has form of *Hotelling's* T^2 *test*

$$T^2 := n(\underline{\overline{X}} - \underline{\mu}_o)^T \mathsf{S}^{-1}(\underline{\overline{X}} - \underline{\mu}_o) \overset{H_1}{\underset{H_0}{\gtrless}} \gamma,$$

where S is the (unbiased) sample covariance

$$\mathsf{S} = \frac{1}{n-1} \sum_{k=1}^{n} (\underline{X}_k - \underline{\overline{X}})(\underline{X}_k - \underline{\overline{X}})^T.$$

We use the following result to set the threshold of the GLRT

FACT: Under H_0, Hotelling's T^2 is distributed as a $\mathcal{T}^2$ distributed r.v. with $(p, n-p)$ d.f. [50, 57].

Thus the level α GLRT of (134) is

$$T^2 := n(\underline{\overline{X}} - \underline{\mu}_o)^T \mathsf{S}^{-1}(\underline{\overline{X}} - \underline{\mu}_o) \underset{H_0}{\overset{H_1}{\gtrless}} \mathcal{T}^{-2}_{p,n-p}(1-\alpha)$$

REMARKS

1. The Hotelling T^2 test is CFAR since under H_0 its distribution is independent of $\mathbf{R}$
2. The T^2 statistic is equal to a F-statistic within a scale factor

$$T^2 = \frac{p(n-1)}{n-p} F_{p,n-p}$$

3. An equivalent test is therefore

$$T^2 := n(\underline{\overline{X}} - \underline{\mu}_o)^T \mathsf{S}^{-1}(\underline{\overline{X}} - \underline{\mu}_o) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{p(n-1)}{n-p} F^{-1}_{p,n-p}(1-\alpha).$$

12.3 TEST OF EQUALITY OF TWO MEAN VECTORS

Assume that we are given two i.i.d. vector samples

$\mathbf{X} = [\underline{X}_1, \ldots, \underline{X}_{n_1}], \quad \underline{X}_i \sim \mathcal{N}_p(\underline{\mu}_x, \mathbf{R})$

$\mathbf{Y} = [\underline{Y}_1, \ldots, \underline{Y}_{n_2}], \quad \underline{Y}_i \sim \mathcal{N}_p(\underline{\mu}_y, \mathbf{R})$

where $n_1 + n_2 = n$. Assume that these samples have the same covariance matrix $\mathbf{R}$ but possibly different means $\underline{\mu}_x$ and $\underline{m}_y$, respectively. It is frequently of interest to test equality of these two means

$$\begin{aligned} H_0 : \underline{\mu}_x - \underline{\mu}_y &= \underline{\Delta}, \ \ \mathbf{R} > 0 \\ H_1 : \underline{\mu}_x - \underline{\mu}_y &\neq \underline{\Delta}, \ \ \mathbf{R} > 0. \end{aligned}$$

The derivation of the GLRT for these hypotheses is simple when inspired by elements of our previous derivation of the GLRT for double sided tests on means of two scalar populations (Sec. 9.5). The GLRT is

$$\sqrt{\frac{n_1 n_2}{n}} (\underline{\overline{Y}} - \underline{\overline{X}} - \underline{\Delta})^T \mathsf{S}_2^{-1} (\underline{\overline{Y}} - \underline{\overline{X}} - \underline{\Delta}) \underset{H_0}{\overset{H_1}{\gtrless}} \mathcal{T}^{-2}_{p,n-p-1}(1-\alpha) \tag{135}$$

where we have defined the pooled sample covariance

$$\mathsf{S}_2 = \frac{1}{n-2} \left(\sum_{i=1}^{n_1} (\underline{X}_i - \underline{\hat{\mu}})(\underline{X}_i - \underline{\hat{\mu}})^T + \sum_{i=1}^{n_2} (\underline{Y}_i - \underline{\hat{\mu}})(\underline{Y}_i - \underline{\hat{\mu}})^T \right),$$

and $\underline{\hat{\mu}} = \frac{1}{2n} \sum_{i=1}^{n} (\underline{X}_i + \underline{Y}_i)$. In analogy to the paired t-test of Sec. 9.5, the test (135) is called the *multivariate paired t-test.*

12.4 TEST OF INDEPENDENCE

n i.i.d. vector samples

$\mathbf{X} = [\underline{X}_1, \ldots, \underline{X}_n]$, $\underline{X}_i \sim \mathcal{N}_p(\underline{\mu}, \mathbf{R})$

To test

$$\begin{aligned} H_0 : \mathbf{R} &= \mathrm{diag}(\sigma_j^2) \\ H_1 : \mathbf{R} &\neq \mathrm{diag}(\sigma_j^2) \end{aligned}$$

with mean vector $\underline{\mu}$ unknown

$$\Lambda_{\mathrm{GLR}} = \frac{\max_{\mathbf{R}\neq\mathrm{diag},\underline{\mu}} f(\mathbf{X}; \underline{\mu}, \mathbf{R})}{\max_{\mathbf{R}=\mathrm{diag},\underline{\mu}} f(\mathbf{X}; \underline{\mu}, \mathbf{R})} =$$

$$\frac{\max_{\mathbf{R}>0} \ |\mathbf{R}|^{-n/2} \ \exp\left(-\frac{1}{2} \sum_{k=1}^n (\underline{X}_k - \overline{\underline{X}})^T \mathbf{R}^{-1} (\underline{X}_k - \overline{\underline{X}})\right)}{\max_{\sigma_j^2>0} \ \left(\prod_{k=1}^p \sigma_k^2\right)^{-n/2} \ \exp\left(-\frac{1}{2} \sum_{k=1}^n \frac{1}{\sigma_k^2} \|\underline{X}_k - \overline{\underline{X}}\|^2\right)}$$

Using previous results

$$\Lambda_{\mathrm{GLR}} = \left(\frac{\prod_{j=1}^p \hat{\sigma}_j^2}{|\hat{\mathbf{R}}|}\right)^{n/2} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \gamma$$

where we have the variance estimate for each channel (row) of $\mathbf{X}$

$$\hat{\sigma}_j^2 := \frac{1}{n} \sum_{k=1}^n (\underline{X}_k - \overline{\underline{X}})_j^2$$

For n sufficiently large we can set the threshold γ using the usual Chi-square asymptotics described in Eq. (113) and discussed in Chapter 8. For this analysis we need calculate the number of degrees of freedom ν of the test statistic under H_0. Recall from that discussion that the degrees of freedom ν is the number of parameters that are unknown under H_1 but are fixed under H_0. We count these parameters as follows For n large we can set γ by using Chi-square asymptotics.

1. $p^2 - p = p(p-1)$ off diagonals in $\mathbf{R}$

2. 1/2 of these off diagonals elements are identical due to symmetry of $\mathbf{R}$

$\Rightarrow \nu = p(p-1)/2$

Thus we obtain the approximate level α GLRT:

$$2 \ln \Lambda_{\mathrm{GLR}} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \gamma' = \chi^{-1}_{p(p-1)/2}(1-\alpha).$$

12.5 TEST OF WHITENESS

n i.i.d. vector samples

$\mathbf{X} = [\underline{X}_1, \ldots, \underline{X}_n]$, $\underline{X}_i \sim \mathcal{N}_p(\underline{\mu}, \mathbf{R})$

To test

$$\begin{aligned} H_0 : \mathbf{R} &= \sigma^2 \mathbf{I} \\ H_1 : \mathbf{R} &\neq \sigma^2 \mathbf{I} \end{aligned}$$

with mean vector $\underline{\mu}$ unknown

$$\begin{aligned} \Lambda_{\text{GLR}} &= \frac{\max_{\mathbf{R} \neq \sigma^2 \mathbf{I}, \underline{\mu}} f(\mathbf{X}; \underline{\mu}, \mathbf{R})}{\max_{\mathbf{R} = \sigma^2 \mathbf{I}, \underline{\mu}} f(\mathbf{X}; \underline{\mu}, \mathbf{R})} \\ &= \frac{\max_{\mathbf{R} > 0} \; |\mathbf{R}|^{-n/2} \; \exp\left(-\frac{1}{2} \; \sum_{k=1}^n (\underline{X}_k - \overline{\underline{X}})^T \mathbf{R}^{-1} (\underline{X}_k - \overline{\underline{X}})\right)}{\max_{\sigma^2 > 0} \; (\sigma^{2p})^{-n/2} \; \exp\left(-\frac{1}{2\sigma^2} \; \sum_{k=1}^n \|\underline{X}_k - \overline{\underline{X}}\|^2\right)} \end{aligned}$$

Or we have (similarly to before)

$$\Lambda_{\text{GLR}} = \left(\frac{\hat{\sigma}^{2p}}{|\hat{\mathbf{R}}|}\right)^{n/2} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \gamma \tag{136}$$

where

$$\begin{aligned} \hat{\sigma}^2 &:= \frac{1}{np} \underbrace{\sum_{k=1}^n \|\underline{X}_k - \overline{\underline{X}}\|^2}_{n \text{trace}\{\hat{\mathbf{R}}\}} \\ &= \frac{1}{p} \text{trace}\left\{\hat{\mathbf{R}}\right\} \end{aligned}$$

and we have defined the covariance estimate

$$\hat{\mathbf{R}} := \frac{1}{n} (\mathbf{X} - \overline{\underline{X}} 1^T)(\mathbf{X} - \overline{\underline{X}} 1^T)^T$$

The GLRT (136) can be represented as a test of the ratio of arithmetic mean to geometric mean of the eigenvalues of the covariance estimate:

$$(\Lambda_{\text{GLR}})^{2/(np)} = \frac{\hat{\sigma}^2}{|\hat{\mathbf{R}}|^{1/p}} \tag{137}$$

$$= \frac{p^{-1}\sum_{i=1}^p \lambda_i^{\hat{\mathbf{R}}}}{\prod_{i=1}^p \left(\lambda_i^{\hat{\mathbf{R}}}\right)^{1/p}} \quad \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} \quad \gamma. \tag{138}$$

With this form we have the interpretation that the GLRT compares the elliptical contour of the level set of the sample's density under H_1 to the spherical contour of the level sets of the sample's density under H_0. The GLRTs (136) and (138) are CFAR tests since the test statistics do not depend on the mean $\underline{\mu}$ or the variance σ^2 of the sample.

PERFORMANCE OF GLRT

For n sufficiently large we again set the threshold γ using the usual Chi-square asymptotics described in Eq. (113) of Chapter 8. We must calculate the number of degrees of freedom ν of the test statistic under H_0: ν being the number of parameters that are unknown under H_1 but that are fixed under H_0. We count these parameters as follows

1. $p(p-1)/2$ elements in the triangle above the diagonal of $\mathbf{R}$ are unknown under H_1 but zero under H_0

2. $p-1$ parameters on the diagonal of $\mathbf{R}$ are unknown under H_1 but known (equal to the common parameter σ^2)under H_0.

We therefore conclude that $\nu = p(p-1)/2 + p - 1 = p(p+1)/2 - 1$ and therefore the Chi square approximation specifies the GLRT with approximate level α as

$$2 \ln \Lambda_{\mathrm{GLR}} \quad \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} \quad \gamma' = \chi^{-1}_{p(p+1)/2-1}(1-\alpha)$$

12.6 CONFIDENCE REGIONS ON VECTOR MEAN

Recall: from the level α double sided test of vector mean we know

$$P_{\underline{\theta}}\left(n(\overline{\underline{X}} - \underline{\mu}_o)^T \mathsf{S}^{-1}(\overline{\underline{X}} - \underline{\mu}_o) \;>\; \mathcal{T}^{-2}_{p,n-p}(1-\alpha)\right) = \alpha$$

where $\underline{\theta} = [\underline{\mu}, \mathbf{R}]$.

Equivalently

$$P_{\underline{\theta}}\left(n(\overline{\underline{X}} - \underline{\mu}_o)^T \mathsf{S}^{-1}(\overline{\underline{X}} - \underline{\mu}_o) \;\leq\; \mathcal{T}^{-2}_{p,n-p}(1-\alpha)\right) = 1-\alpha$$

This is a "simultaneous confidence statement" on all elements of mean vector $\underline{\mu}$ for unknown covariance $\mathbf{R}$ given measurement $\mathbf{X}$

$\Rightarrow$ $(1-\alpha)\%$ confidence region on $\underline{\mu}$ is the ellipsoid

$$\left\{\underline{\mu} : n(\overline{\underline{X}} - \underline{\mu})^T \mathsf{S}^{-1}(\overline{\underline{X}} - \underline{\mu}) \;\leq\; \mathcal{T}^{-2}_{p,n-p}(1-\alpha)\right\}$$

.

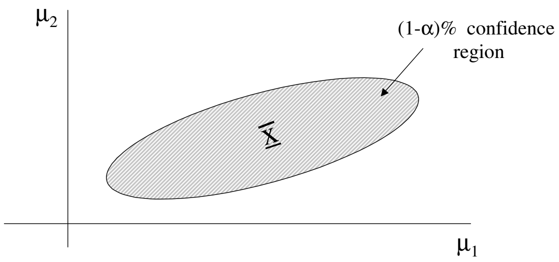


Figure 183: *Confidence region for all elements of mean vector* $\underline{\mu}$ *is an ellipsoid*

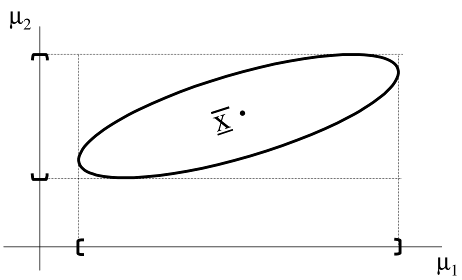


Figure 184: *Confidence ellipsoid gives "marginal" confidence intervals on each element of* $\underline{\mu} = [\mu_1, \ldots, \mu_p]^T$

12.7 EXAMPLES

Example 50 *Confidence band on a periodic signal in noise*

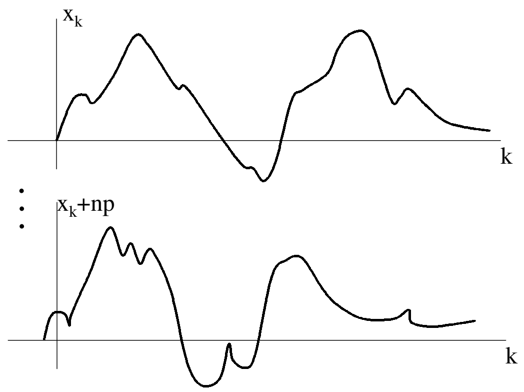


Figure 185: *Multiple uncorrelated measurements of a segment of a periodic signal.*

$$x_k = s_k + v_k$$

* $s_k = s_{k+nT_p}$: unknown periodic signal with known period T_p

* v_k: zero mean w.s.s. noise of bandwidth $1/(MT_p)$ Hz

Step 1: construct measurement matrix

$$\underline{X}_i = [x_{1+(i-1)MT_p}, \ldots, x_{T_p+(i-1)MT_p}]^T$$

Step 2: find conf. intervals on each s_k from ellipsoid

$$[(\overline{\underline{X}})_k - l_k \ \leq s_k \leq (\overline{\underline{X}})_k + u_k]$$

Example 51 *CFAR signal detection in narrowband uncalibrated array*

k-th snapshot of p-sensor array output:

$$\underline{x}_k = \underline{a}\, s + \underline{v}_k, \quad k = 1, \ldots, n. \tag{139}$$

* $\underline{a}$: unknown array response (steering) vector

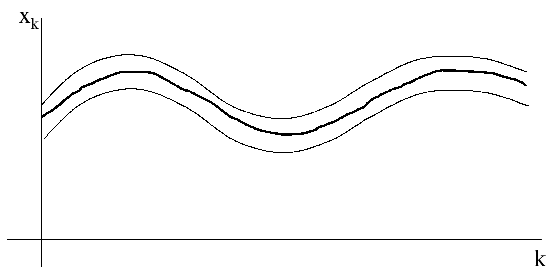


Figure 186: *Confidence band on signal over one signal period.*

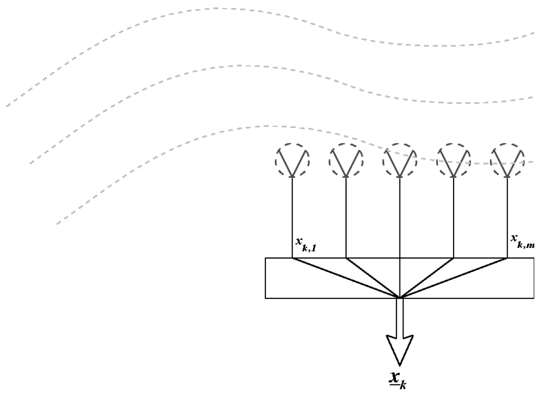


Figure 187: *Sensor array generates spatio-temporal measurement.*

* $\underline{v}_k$: Gaussian $\mathcal{N}_p(0, \mathbf{R})$ array noise vector with unknown spatial covariance $\mathbf{R}$

* s: unknown deterministic signal amplitude

Objective: detect presence of any non-zero signal amplitude at level α

$$H_0 : s = 0, \quad k = 1, \ldots, n$$

$$H_1 : s \neq 0, \quad k = 1, \ldots, n$$

This is equivalent to

$$H_0 : E[\underline{X}_i] = \underline{\mu} = 0, \;\; \mathbf{R} > 0$$

$$H_1 : E[\underline{X}_i] = \underline{\mu} \neq 0, \;\; \mathbf{R} > 0$$

For which we know:

* level α GLRT is the Hotelling T^2 test

* confidence region for $\underline{\mu} = \underline{a}s$ is an ellipsoid.

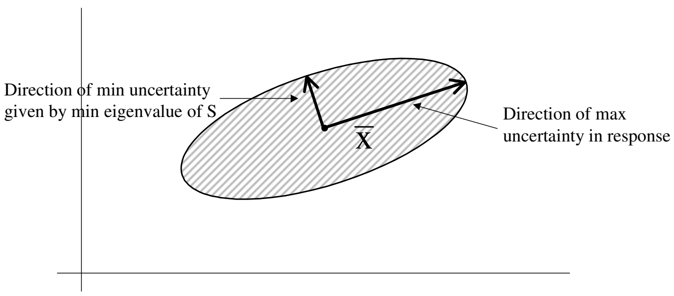


Figure 188: *Confidence region for array response vector $\underline{a}s$ is an ellipse in 2D.*

12.8 BACKGROUND REFERENCES

Many of the GLRT results in this chapter can be found in the books by Morrison [50] and Anderson [2]. Some applications of these results to signal and array processing problems are discussed in Van Trees [73]. More applications of detection theory to multi-channel problems arising in array processing and spectral estimation can be found in Haykin [24] and Stoica and Moses [69]. The books by Eaton [15], Mardia, Kent and Bibby [44], and Muirhead [51] give more advanced treatments of general multivariate analysis techniques and testing of composite hypotheses. The problem of constructing confidence regions for vector parameter is closely related to the problem of simultaneous confidence intervals and this topic is covered in detail by Miller [47]. Miller's book does not cover the popular and more flexible False Discovery Rate (FDR) as an alternative to confidence level, for which the reader is referred to Benjamini and Yekutieli's paper [5] and its hypothesis testing homolog by Benjamini and Hochberg [4].

12.9 EXERCISES

12.1 Extend the multivariate paired-t test derived in Sec. 12.3 to the case where $\underline{x}_i \sim \mathcal{N}(\underline{\mu}_x, R_x)$ and $\underline{y}_i \sim \mathcal{N}(\underline{\mu}_y, R_y)$ for the case that the two covariance matrices R_x and R_y may be unequal and are unknown. How many degrees of freedom does the the asymptotic Chi-square distribution have?

12.2 In Example 51 the optimal CFAR detector for a scalar signal s viewed from a p-sensor array output with array response $\underline{a}$ and noise $\underline{v}_k$. In this problem we extend this to CFAR detection of multiple (m) scalar signals $\underline{s} = [s_1, \ldots, s_m]$ following the observation model:

$$\underline{x}_k = A\underline{s} + \underline{v}_k, \quad k = 1, \ldots, n \tag{140}$$

where $A = [\underline{a}_1, \ldots, \underline{a}_p]$ is an unknown $p \times m$ matrix and $\underline{v}_k$ are i.i.d. $\mathcal{N}(0, R)$ random vectors with unknown covariance R. Derive the GLRT for this problem. How many degrees of freedom does the the asymptotic Chi-square distribution have?

12.3 Consider the same model as (139) but assume that s is a Gaussian distributed random variable and $\underline{a}$ and R are unknown. Derive the GLRT.

12.4 Consider the same scalar model as (139) but now assume that $\underline{a}$ is known while the noise covariance R is unknown. Derive the GLRT.

12.5 Extend the analysis of the previous problem to the multiple signal case (140) when A has columns of sinusoidal form:

$$\underline{a}_k = [1, \cos(2\pi f_k), \ldots, \cos(2\pi f_k(p-1))]^T, \quad k = 1, \ldots, m$$

while the noise covariance R is unknown. Derive the GLRT (you may assume that the $\underline{a}_k$'s are orthogonal if you wish).

End of chapter

13 BIBLIOGRAPHY

References

[1] M. Abromowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover, New York, N.Y., 1965.

[2] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958.

[3] M. Basseville and A. Benveniste, editors, *Detection of abrupt changes in signals and dynamical systems*, Springer Lecture Notes in Control and Information Sciences, 1986.

[4] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Royal Statistical Society*, vol. 57, pp. 289–300, 1995.

[5] Y. Benjamini and D. Yekutieli, "False discovery rate adjusted confidence intervals for selected parameters (preprint)," *J. Am. Statist. Assoc.*, vol. 100, no. 469, pp. 71–81, March 2005.

[6] R. Beran, "Pivoting to reduce level error of confidence sets," *Biometrika*, vol. 74, pp. 457–468, 1987.

[7] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, Holden-Day, San Francisco, 1977.

[8] H. Bode and C. Shannon, "A simplified derivation of linear least squares smoothing and prediction theory," *Proc. of the Institute of Radio Engineers (IRE)*, vol. 38, pp. 417–425, April 1950.

[9] S. Bose and A. O. Steinhardt, "Adaptive array detection of uncertain rank one waveforms," *Signal Processing*, vol. 44, no. 11, pp. 2801–2809, Nov. 1996.

[10] S. Bose and A. O. Steinhardt, "A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices," *Signal Processing*, vol. 43, no. 9, pp. 2164–2175, Sept. 1995.

[11] D. R. Brillinger, *Time Series: Data Analysis and Theory*, Springer-Verlag, New York, 1981.

[12] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer-Verlag, New York, 1987.

[13] W. Davenport and W. Root, *An introduction to the theory of random signals and noise*, IEEE Press, New York (reprint of 1958 McGraw-Hill edition), 1987.

[14] M. DeGroot, *Optimal Statistical decisions*, McGraw Hill, NJ, 1970.

[15] M. L. Eaton, *Multivariate Statistics - A Vector Space Approach*, Wiley, New York, 1983.

[16] T. S. Ferguson, *Mathematical Statistics - A Decision Theoretic Approach*, Academic Press, Orlando FL, 1967.

[17] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman and Hall/CRC, Boca Raton, 1995.

[18] M. Gevers and T. Kailath, "An innovations approach to least squares estimation, part vi: Discrete-time innovations representations and recursive estimation," *IEEE Trans. Automatic Control*, vol. AC-18, pp. 588–600, Dec. 1973.

[19] G. H. Golub and C. F. Van Loan, *Matrix Computations (2nd Edition)*, The Johns Hopkins University Press, Baltimore, 1989.

[20] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. on Inform. Theory*, vol. IT-36, pp. 1285–1301, Nov. 1990.

[21] F. A. Graybill, *Matrices with Applications in Statistics*, Wadsworth Publishing Co., Belmont CA, 1983.

[22] G. Hardy, J. Littlewood, and G. Pólya, *Inequalities (2nd Edition)*, Cambridge Univ. Press, 1951.

[23] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, Wiley, New York, 1996.

[24] S. Haykin, *Array Signal Processing*, Prentice-Hall, Englewood Cliffs NJ, 1985.

[25] C. Helstrom, *Elements of signal detection and estimation*, Prentice-Hall, Englewood Cliffs, 1995.

[26] J. U. Hjorth, *Computer intensive statistical methods*, CHapman and Hall, London, 1994.

[27] D. C. Hoaglin, F. Mosteller, and J. W. Tukey, *Understanding robust and exploratory data analysis*, Wiley, New York, 1983.

[28] M. Hollander and D. A. Wolfe, *Nonparametric statistical methods (2nd Edition)*, Wiley, New York, 1991.

[29] I. A. Ibragimov and R. Z. Has'minskii, *Statistical estimation: Asymptotic theory*, Springer-Verlag, New York, 1981.

[30] N. L. Johnson, S. Kotz, and A. W. N. Balakrishnan, *Continuous univariate distributions: Vol. 1*, Wiley, New York, 1994.

[31] N. L. Johnson, S. Kotz, and A. W. N. Balakrishnan, *Continuous univariate distributions: Vol. 2*, Wiley, New York, 1995.

[32] T. Kailath, *Lectures on Wiener and Kalman Filtering*, Springer-Verlag, New York, 1981.

[33] T. Kariya and B. K. Sinha, *Robustness of Statistical Tests*, Academic Press, San Diego, 1989.

[34] R. Kass and P. Vos, *Geometrical Foundations of Asymptotic Inference*, Wiley, New York, 1997.

[35] S. Kassam and J. Thomas, *Nonparametric detection - theory and applications*, Dowden, Hutchinson and Ross, 1980.

[36] S. M. Kay, *Statistical Estimation*, Prentice-Hall, Englewood-Cliffs N.J., 1991.

[37] H. Kim and A. Hero, "Comparison of GLR and invariant detectors under structured clutter covariance," *IEEE Trans. on Image Processing*, vol. 10, no. 10, pp. 1509–1520, Oct 2001.

[38] G. F. Knoll, *Radiation Detection*, Wiley, New York, 1985.

[39] E. L. Lehmann, *Testing Statistical Hypotheses*, Wiley, New York, 1959.

[40] E. L. Lehmann, *Theory of Point Estimation*, Wiley, New York, 1983.

[41] E. L. Lehmann, *Theory of Point Estimation, 2nd Edition*, Wiley, New York, 1991.

[42] D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1969.

[43] E. B. Manoukian, *Modern Concepts and Theorems of Mathematical Statistics*, Springer-Verlag, New York N.Y., 1986.

[44] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*, Academic Press, New York, 1979.

[45] J. M. Mendel, *Lessons in estimation for signal processing, communications, and control*, Prentice-Hall, Englewood Cliffs NJ, 1995.

[46] D. Middleton, "Statistical methods for the detection of pulsed radar in noise," in *Communication Theory*, W. Jackson, editor, pp. 241–270, Academic Press, New York, 1953.

[47] R. G. Miller, *Simultaneous Statistical Inference*, Springer-Verlag, NY, 1981.

[48] A. M. Mood, F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, McGraw Hill, New York, 1976.

[49] T. K. Moon and W. C. Stirling, *Mathematical methods and algorithms for signal processing*, Prentice Hall, Englewood Cliffs, 2000.

[50] D. F. Morrison, *Multivariate statistical methods*, McGraw Hill, New York, 1967.

[51] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.

[52] B. Noble and J. W. Daniel, *Applied Linear Algebra*, Prentice Hall, Englewood Cliffs, NJ, 1977.

[53] A. V. Oppenheim and A. S. Willsky, *Signals and Systems*, Prentice-Hall, Englewood Cliffs, N.J., 1983.

[54] A. Papoulis, *Probability, random variables, and stochastic processes (3rd ed)*, McGraw Hill, New York, N.Y., 1991.

[55] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 1988.

[56] J. G. Proakis and D. G. Manolakis, *Digital signal processing: principles, algorithms, and applications*, Prentice-Hall, NY, 1996.

[57] C. R. Rao, *Linear Statistical Inference and Its Applications*, Wiley, New York, 1973.

[58] J. P. Romano and A. F. Siegel, *Counterexamples in probability and statistics*, Wadsworth, Belmont CA, 1983.

[59] S. Ross, *A first course in probability*, Prentice-Hall, Englewood Cliffs, N.J., 1998.

[60] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, Reading, MA, 1991.

[61] L. L. Scharf and D. W. Lytle, "Signal detection in Gaussian noise of unknown level: an invariance application," *IEEE Trans. on Inform. Theory*, vol. IT-17, no. 3, pp. 404–411, 1971.

[62] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, Wiley, New York, 1980.

[63] M. Shao and C. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," *Proceedings of the IEEE*, vol. 81, no. 7, pp. 986–1010, July 1993.

[64] G. E. Shilov, *Linear Algebra*, Dover Publications, New York, N.Y., 1977.

[65] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*, Springer-Verlag, New York, 1991.

[66] T. Soderstrom and P. Stoica, *System identification*, Prentice Hall, Englewood Cliffs NJ, 1989.

[67] M. D. Srinath, P. K. Rajasekaran, and R. Viswanathan, *Introduction to statistical signal processing with applications*, Prentice Hall, Englewood Cliffs, 1996.

[68] H. Stark and J. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, Prentice-Hall, Englewood Cliffs, N.J., 1986.

[69] P. Stoica and R. Moses, *Introduction to spectral analysis*, Prentice-Hall, Englewood Cliffs NJ, 1997.

[70] P. Strobach, *Linear Prediction Theory: A Mathematical Basis for Adaptive Systems*, Springer-Verlag, New York, 1990.

[71] M. A. Tanner, *Tools for Statistical Inference; Methods for the exploration of posterior distributions and likelihood functions*, Springer-Verlag, New York, 1993.

[72] J. B. Thomas, *An Introduction to Statistical Communications Theory*, Wiley, New York, 1969.

[73] H. L. Van-Trees, *Detection, Estimation, and Modulation Theory: Part I*, Wiley, New York, 1968.

[74] H. L. Van-Trees, *Detection, Estimation, and Modulation Theory: Part III*, Wiley, New York, 2001.

[75] R. E. Walpole, R. H. Myers, and S. L. Myers, *Probability and statistics for engineers and scientists*, Prentice Hall, New Jersey, 1998.

[76] A. D. Whalen, *Detection of Signals in Noise (2nd Ed.)*, Academic Press, Orlando, 1995.